

NASA Contractor Report 4564  
DOT/FAA/RD-93/41

# On the Estimation of the Correlation Dimension and Its Application to Radar Reflector Discrimination

Kevin D. Barnett  
*Clemson University*  
*Clemson, South Carolina*

Prepared for  
Langley Research Center  
under Grant NAG1-928



National Aeronautics and  
Space Administration

Office of Management

Scientific and Technical  
Information Program

1993

N94-23304

Unclass

H1/33 0203381

(NASA-CR-4564) ON THE ESTIMATION  
OF THE CORRELATION DIMENSION AND  
ITS APPLICATION TO RADAR REFLECTOR  
DISCRIMINATION Final Report  
(Clemson Univ.) 72 p



## TABLE OF CONTENTS

	Page
Table of Contents .....	iii
List of Figures .....	v
Acknowledgements .....	vii
CHAPTER	
1. INTRODUCTION .....	1
Chaotic Systems and the Correlation Dimension .....	1
Review of Previous Work in Correlation Dimension	
Estimation .....	2
Contribution of This Work .....	3
Organization of the Thesis .....	4
2. DYNAMICAL SYSTEMS: MATHEMATICAL AND GEOMETRICAL CONCEPTS .....	6
Theory of Dynamical Systems .....	6
Mathematical Model of a Dynamical System .....	6
The Attractor of a Dynamical System .....	7
Chaotic Dynamical Systems .....	8
Nature of Chaotic Attractors .....	9
Examples of Chaotic Systems .....	12
3. THE CORRELATION DIMENSION .....	17
Measures of Dimension .....	17
The Correlation Integral and Correlation Dimension .....	19
Definition of Correlation Integral .....	19
Shape of Correlation Integral .....	19
Reconstruction of System Dynamics from a Time Series .....	22
The Number of Independent Measurements Required .....	22
The Method of Time Delay Coordinates .....	23

## Table of Contents (Continued)

	Page
Relationship between Time Delay and Phase Space	
Reconstructions .....	24
Influence of Time Delay on Coordinates' Independence .....	24
A Conjecture on the Appropriateness of Uniform Time	
Delay Based on Mutual Information .....	30
4. A NEW ESTIMATE FOR THE CORRELATION	
DIMENSION .....	33
Traditional Method Using Least Squares .....	33
A New Method .....	34
Error Criterion and Justification .....	34
Advantages over Other Methods .....	35
Comparison of Results: New Method Versus Least Squares .....	36
5. APPLICATION OF THE CORRELATION DIMENSION TO	
RADAR REFLECTOR DISCRIMINATION .....	48
Radar System Scenario .....	48
Method of Analysis .....	49
Results of Correlation Dimension Analysis .....	50
6. CONCLUSIONS .....	58
Discussion of Results .....	58
New Estimate of the Correlation Dimension .....	58
Correlation Dimension as a Radar Reflector Discriminant .....	59
Conclusions and Future Work .....	60
REFERENCES .....	62

## LIST OF FIGURES

Figure	Page
2.1 Koch curve as an example of a curve with fractional dimension. ....	11
2.2 Sample attractor for the logistic equation with $\mu = 3.5699456$ and initial condition $x[0] = 0.6791$ . ....	14
2.3 Sample Lorenz attractor simulated using fourth order Runge-Kutta numerical integration scheme. ....	15
2.4 Sample Rössler attractor with parameter values $a = 0.15$ , $b = 0.20$ , and $c = 10.0$ . ....	16
3.1 Correlation integral for linear attractor with equally spaced points. ....	20
3.2 Estimates of the correlation dimension for linear attractor with equally spaced points. ....	21
3.3 Reconstruction of Lorenz attractor using time delay coordinates with time delay too small, $T = 1$ in this case. ....	25
3.4 Reconstruction of Lorenz attractor using time delay coordinates with time delay too large, $T = 77$ in this case. ....	27
3.5 Mutual information between delayed versions of the x component of the Lorenz system. ....	28
3.6 Reconstruction of Lorenz attractor using time delay coordinates with time delay chosen as the first minimum of the mutual information function, $T = 17$ . ....	29
3.7 Mutual information between delayed versions of the x component of the Rössler system. ....	31
4.1 Assumed form for a typical $\log[C(r)]$ vs. $\log[r]$ curve. ....	37
4.2 Correlation integral for the logistic equation. ....	39

## List of Figures (Continued)

	Page
4.3 Estimates of the correlation dimension of the logistic equation using both the new method and least squares. ....	40
4.4 Correlation integral for the Lorenz system. ....	41
4.5 Estimates of the correlation dimension of the Lorenz system. ....	42
4.6 Correlation integral for the Gaussian noise sequence. ....	44
4.7 Estimates of the correlation dimension for the Gaussian noise sequence using both the new method and least squares. ....	45
4.8 Correlation integral for the Lorenz system with zero mean, unit variance Gaussian noise added. ....	46
4.9 Estimates of the correlation dimension for the Lorenz system with zero mean, unit variance Gaussian noise added using both the new estimate and least squares. ....	47
5.1 Calculated correlation dimension for ground clutter for the first test. ....	51
5.2 Calculated correlation dimension for weather for the first test. ....	52
5.3 Calculated correlation dimension for ground clutter for the second test. ....	53
5.4 Calculated correlation dimension for weather for the second test. ....	54
5.5 Density functions of correlation dimension values conditioned on ground returns and weather returns for the first test. ....	55
5.6 Density functions of correlation dimension values conditioned on ground returns and weather returns for the second test. ....	56

## ACKNOWLEDGMENTS

The author would like to express his appreciation to all those who contributed in some manner to this work. In particular, he would like to thank his advisor, Dr. Ernest G. Baxa, Jr., for his guidance, and the Ph.D. committee members, Dr. John Komo, Dr. Clark Jeffries, and Dr. John Gowdy. Their support and helpful criticism are greatly appreciated.

Thanks are also due to NASA Langley Research Center for funding and provision of data and equipment for this work.

Finally, the author would like to thank his family for their encouragement, support, and, most of all, patience. And lastly, thanks to the multitude of students and friends who have been tremendously encouraging during his term at Clemson University; he is very indebted for their support and confidence that this work should and would be completed.





# CHAPTER 1

## INTRODUCTION

### 1.1 Chaotic Systems and the Correlation Dimension

In the past, the vast majority of systems theory work has concentrated on linear systems. However, in recent years the chaotic system model has been found appropriate for several physical systems of interest. The chaotic system is a particular type of nonlinear system which shows an extreme sensitivity to its initial conditions. This characteristic implies that long term prediction of chaotic systems is impossible, but there are other reasons for our interest in these models. One attribute of particular interest in time series analysis is that chaotic time sequences often appear random.

With regards to linear systems, any component which cannot be predicted in time sequences in the past has been attributed to “random noise”, which may be a generic term for that which cannot be modelled with sufficient accuracy. In many cases it would be very beneficial to distinguish between that which is truly random and that which results from chaos, often to know simply what the order of an appropriate model is: high order if true randomness is involved, low order if the system is chaotic. Recent work has shown that this can be done using a measure called the correlation dimension.

Dissipative dynamical systems are characterized by an attractor in phase space which shrinks to zero in volume as time passes. For dissipative chaotic systems, this attractor is also characterized by having fractional dimension. The correlation dimension is a generalized dimension measure which describes the geometric shape of the system’s attractor, weighted according to the frequency of visitation of the system’s state to each portion of the attractor.

## 1.2 Review of Previous Work in Correlation Dimension Estimation

Chaos theory itself was founded with the pioneering work of Lorenz [1], who noted from atmospheric simulations that his nonlinear models were extremely sensitive to perturbations in their initial states. Lorenz termed this phenomenon “deterministic nonperiodic flow” to emphasize the lack of any randomness or periodicity in the system.

The estimation of attractor dimension was relatively undeveloped until work in the early 1980s [2] began to focus on the dimension of the attractor of chaotic systems as a physically meaningful attribute. Also, the work of Packard *et al.* [3] showed that it is possible to derive the geometry of attractors from a single time series using time delay embedding, which is important since vectors are rarely available.

Much work in estimation of the dimension of chaotic attractors ensued, the most notable being the works of Grassberger and Procaccia [4, 5] and Farmer *et al.* [6]. Researchers began to recognize that the strictly geometric dimension of an attractor is not physically as significant as are measures which expose the dynamics on the attractor as well as the geometry. In this latter category is the correlation dimension, introduced by Grassberger and Procaccia [4, 5]. This measure has the advantages that it is easy to understand, easy to calculate, and contains information about both the attractor geometry and the system’s dynamics on the attractor.

These developments made it possible to derive a physically meaningful invariant of a system directly from a single time sequence by finding the slope of a log – log plot of the correlation integral. The question of what embedding dimension to use with the time delay embedding procedure was answered formally by application of the Takens Embedding Theorem [7]. Takens developed an optimal estimator for the correlation dimension [8]. A criterion for the choice of time delay using the mutual information function was introduced by Fraser and Swinney [9].

Later, the extensive works of Theiler [10, 11, 12] exposed many of the pitfalls and gave some suggestions for overcoming problems with noise, autocorrelation in

data, lacunarity, and suggested a new method for the Takens' estimator calculated directly from the correlation integral. More recently, chaotic theory and correlation dimension estimation have been adopted and applied in the realm of signal processing, in applications such as speech and radar [13, 14, 15, 16, 17, 18].

However, many details are still left to be determined. In particular, it is well known [11] that only a certain region of the correlation integral can be used for estimation of the correlation dimension, but there are as yet no firm guidelines for choosing that region. Further, theoretically the line fitting to the log – log plot of the correlation integral has been argued to be unjustified [11].

Empirical evidence about the accuracy and discriminatory power of dimension estimation has been mixed. On the one hand, it has been shown very effective in discriminating between chaos and noise, while on the other hand many researchers have questioned the validity of terming the estimates actual “dimensions”.

### 1.3 Contribution of This Work

Two contributions are presented here: First the problem of estimating the correlation dimension from the correlation integral is attacked. As previously noted, un-weighted least squares has been criticized for not being theoretically justified. Meanwhile, the optimal estimator of Takens does not show significantly better results, though it does provide a meaningful estimate of its error along with the dimension estimate. Results using the weighted least squares methods do not generally seem more promising either, in addition to the problems in choosing “appropriate” weights.

Here a new method is proposed for calculating the correlation dimension from the correlation integral. The new method differs from least squares in the error criterion chosen: since there is much statistical dependence between the points of the correlation integral, the error criterion is based on the error in the *increments* of the correlation integral. Further, a measure based on the increments tends to not be influenced when the correlation integral becomes flat, so there is no need to specify a point at which to stop on the correlation integral.

Second, as an application, the correlation dimension is used in the discrimination between radar returns from weather and ground clutter [19]. This application is directed towards an ongoing effort to detect low altitude windshear in the near terminal area of airports, as described by Fujita [20], Bracalente *et al.* [21], and Baxa [22]. This presents a plethora of problems for conventional signal processing, and since weather has been claimed by Lorenz to be a low dimensional chaotic system [1], this low dimension should be reflected in the correlation dimension of the radar return.

## 1.4 Organization of the Thesis

The study of correlation dimension estimation begins with a discussion of dynamical systems theory in Chapter 2, followed by an introduction to chaotic dynamical systems.

The notion of fractional dimensions is introduced in Chapter 3. Several dimensions are considered, starting with the purely geometrical Hausdorff dimension and continuing to the more information theoretic correlation dimension. The correlation dimension is then defined more formally in terms of the correlation integral. Chapter 3 continues with some practical considerations in estimating the correlation integral and the correlation dimension and concludes with a conjecture on the appropriateness of the mutual information criterion in the choice of time delay.

Chapter 4 begins with a more detailed discussion of the problems associated with deriving the correlation dimension from the correlation integral. Then a new error criterion for the linear curve fitting is introduced. This error criterion is shown to be more justified statistically than the criterion used with least squares. The new method is also shown to be more easy to apply, since it saturates beyond the linear region of the log – log correlation integral plot. Results using the new method are then compared with those using least squares for several simulated attractors for which the true correlation dimension is known. The new method is shown to give very accurate and consistent predictions of the correlation dimension.

In Chapter 5, the correlation dimension is shown to be effective in several examples of discriminating between actual radar returns from weather and from the ground. The approach is to analyze radar returns which are known to result from weather and others which result from the ground. Several spatially overlapping scans are used and the results averaged to reduce errors in the estimates, since the return sequences are very short.

Conclusions and suggestions for further work are presented in Chapter 6. The overall effectiveness of the new technique for calculating the correlation dimension is discussed, as well as the promise of the correlation dimension applied to discrimination between different radar reflectors.

## CHAPTER 2

### DYNAMICAL SYSTEMS: MATHEMATICAL AND GEOMETRICAL CONCEPTS

#### 2.1 Theory of Dynamical Systems

##### 2.1.1 Mathematical Model of a Dynamical System

Chaotic systems being a special class of dynamical systems, the discussion will begin with an explanation of some terms and concepts from dynamical systems theory. Basic to a dynamical system is the concept of the *order* of the system,  $n$ , which counts the number of active degrees of freedom of the system. This can also be viewed as the dimension of the phase space of the system, or the number of independent quantities required to uniquely specify the *state* of the system at any given time. Phase space is then the  $n$  dimensional hyperspace  $\mathfrak{R}^n$  in which the system's *state vector* moves. Each point in this hyperspace describes a particular state of the system.

The *state* of the system,  $\mathbf{x}(t)$ , is represented by an  $n$  dimensional real vector which gives a complete, unique, description of the system at time  $t$ . The state evolves with time according to the *dynamics* of the particular system. The dynamics are described by the function  $\phi : \mathfrak{R}^n \rightarrow \mathfrak{R}^n$ , where  $\phi$  is a  $C^1$  (continuous with continuous first derivative) function or mapping [23] with the properties:

$$\begin{aligned}\phi_0(\mathbf{x}) &= \mathbf{x} \\ \phi_1(\mathbf{x}) &= \phi(\mathbf{x}) \\ \phi_{t_1+t_2}(\mathbf{x}) &= \phi_{t_1}(\phi_{t_2}(\mathbf{x})).\end{aligned}$$

The preceding discussion applies also to discrete time systems, and since time sequences are more often the concern than continuous time signals, it is worthwhile to consider the discrete time equivalents. The state of a system is now discretized

in time, being denoted by  $\mathbf{x}[k]$ , where  $k$  is the discrete time variable, or time index. The dynamics are given by the function  $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , where  $f$  is a  $C^1$  function often referred to as the *return map*:

$$\mathbf{x}[k + 1] = f\{\mathbf{x}[k]\}.$$

In the discrete time case, the time evolution can be given as a map iterated from the initial state  $\mathbf{x}[0]$ :  $\mathbf{x}[1] = f\{\mathbf{x}[0]\}$ ,  $\mathbf{x}[2] = f\{\mathbf{x}[1]\} = f^2\{\mathbf{x}[0]\}$ , .... The trajectory of the system is then the sequence of points in phase space ( $\mathbb{R}^n$ ) given by  $\mathbf{x}[0]$ ,  $\mathbf{x}[1]$ ,  $\mathbf{x}[2]$ , ....

The observed *output* of a dynamical system is some function of the system's state. It may be one component of the state or a combination of several components, but it is practically always of lower dimension than the state itself; in fact, it is most often a scalar sequence. (Note that observations of topological dimension higher than the order of the system must contain redundant information.) Mathematically, the observation function is then a *projection* from the  $n$  dimensional phase space into a lower dimension, i.e., the output  $\mathbf{y}(t)$  is related to the state  $\mathbf{x}(t)$  by

$$\mathbf{y}(t) = \mathbf{g}(\mathbf{x}(t)),$$

where the function  $\mathbf{g}(\cdot)$  is a mapping from  $n$  dimensional space to an  $m$  dimensional space, with  $n \geq m$ ,  $\mathbf{g} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ . (For the case in which the observation is a scalar sequence,  $m = 1$ .)

### 2.1.2 The Attractor of a Dynamical System

A trajectory is the sequence of points  $\mathbf{x}[0]$ ,  $\mathbf{x}[1]$ , ... for a discrete time system, or the continuous evolution  $\mathbf{x}(t)$ ,  $t \geq 0$  for a continuous time system. The trajectory has two components: a transient part, which the state vector follows initially, and a steady state part, to which the system's trajectory eventually settles.

After a sufficiently long time ("sufficiently long" being determined in general by the system's dynamics), the system state traces out what is termed an *attractor* in

$\mathbb{R}^n$ . For systems which have multiple attractors, the *specific* attractor a system follows is determined by the system's initial state. In particular, the initial condition will be in the *basin of attraction* of the attractor to which the system settles. Formally, Devaney [24] defines an attractor as "...an invariant set to which all nearby orbits [trajectories] converge."

The *Lyapunov exponents* [25] describe the average rates of contraction and expansion with time of the trajectory in different directions in phase space. Specifically, the expansion in direction  $i$  is proportional to  $e^{\lambda_i t}$ , where  $\lambda_i$  is the  $i^{\text{th}}$  Lyapunov exponent; obviously, this entails expansion for  $\lambda_i > 0$  and contraction for  $\lambda_i < 0$ . Further, the sum of the Lyapunov exponents determines the overall expansion or contraction of a reference chunk of phase space. For dissipative dynamical systems, the Lyapunov exponents are all negative, implying that a reference volume in phase space contracts in each direction as time passes. Thus the volume of the part of the attractor also shrinks with time.

## 2.2 Chaotic Dynamical Systems

It has been found that there exists a class of nonlinear systems whose responses are such that they *amplify* differences in initial conditions, rather than forgetting the initial state. This phenomenon was first recorded by Lorenz in 1963 [1], and has since been termed "chaos". There is particular interest in these systems in part due to the fact that sequences associated with chaotic systems appear noiselike though the systems which generate them are completely deterministic. A great deal of effort is being expended in finding measures which would distinguish between chaos and noise, and that is a primary focus of this work.

Chaotic systems exhibit several interesting features, but it is the *sensitive dependence on initial conditions* which distinguishes chaotic and nonchaotic systems [23]. This means that even if the dynamics of the system are known perfectly, with the limited precision available in our world, the initial conditions cannot be specified



precisely enough that the state of the system can be predicted in the long term. This is in contrast to traditional systems theory, in which unpredictability is attributed to the presence of “random noise” in the system. In chaotic systems, on the other hand, the unpredictability is a direct result of the extreme sensitivity to initial conditions.

For initial conditions which are within a given attractor’s basin of attraction, a dissipative linear dynamical system tends to forget its initial state as time passes, all initial conditions within the given basin of attraction converging to the same attractor; the steady state response then is a function of the system’s dynamics alone. By contrast, a chaotic dynamical system tends to amplify differences in initial conditions as time passes; imperceptible differences in the initial conditions grow exponentially as time passes and the difference in the initial conditions becomes distinguishable [23]. The chaotic system can thus be considered an information source [25].

### 2.2.1 Nature of Chaotic Attractors

The sensitivity to initial conditions is related to the Lyapunov exponents of the system. A chaotic system possesses at least one Lyapunov exponent which is positive. (However, the sum of the Lyapunov exponents may be negative, so that the system is nevertheless dissipative.) This means that there is a stretching with time in phase space in at least one direction. As discussed earlier, a dissipative dynamical system (whose Lyapunov exponents are all negative) is characterized by an attractor volume in phase space which shrinks to zero as time evolves. In this case, the attractor becomes a fixed point. By contrast, the attractor of a chaotic system contracts in some directions (those directions corresponding to negative Lyapunov exponents) as it expands in other directions (those directions corresponding to positive Lyapunov exponents), such that the volume tends to zero, but the attractor is not a fixed point [11].

Such an attractor has a dimension which is noninteger and has been termed a *strange attractor* [6, 4, 5]. (Devaney [24] points out that the term “strange attractor”

is not appropriate since most of these attractors have been analyzed; he argues that “hyperbolic attractor” is more appropriate. To be precise, conservative dynamical systems may exhibit chaos, but only dissipative dynamical systems have strange attractors [11].) Further, this is not the only definition of a strange attractor; Eckmann and Ruelle [23] define a strange attractor as an attractor with sensitive dependence on initial conditions: “...the notion of strangeness refers thus to the dynamics on the attractor, and not just to its geometry ...”.

The idea of a *fractional* dimension requires a precise mathematical definition of dimension. The most common concept of dimension corresponds to what is precisely termed the *topological* dimension, and is indeed integer. Geometrically, this dimension corresponds to the smallest number of components in a vector which uniquely defines a point in the space. For example, the dimension of a line is obviously one, since only a single component is required to uniquely describe a point on the line. Further, the dimension is exactly one since a line perfectly fills the one dimension. Thus is also the case of a disc, which has dimension exactly two. However, this dimension may not be sufficient to accurately describe the shape of more complicated geometries, in particular the attractors of dissipative chaotic systems.

Some objects do not entirely fill the space in which they reside. In other words, though it may require 3 components to uniquely describe a point on a given surface, the surface does not fully explore 3-space, but is confined to some well defined subset of the space. In this case the topological dimension is three, but the *Hausdorff* dimension is noninteger and less than three. This more mathematical definition of dimension is best explained by a famous example.

The Koch curve (see Figure 2.1) is a simple example which shows the meaning of “fractional dimension”. This curve is generated by first starting with a line segment of unit length. The middle  $\frac{1}{3}$  segment is replaced by two segments in a sort of “tent” shape, each of length  $\frac{1}{3}$ , giving a curve of length  $\frac{4}{3}$  with the same starting and ending

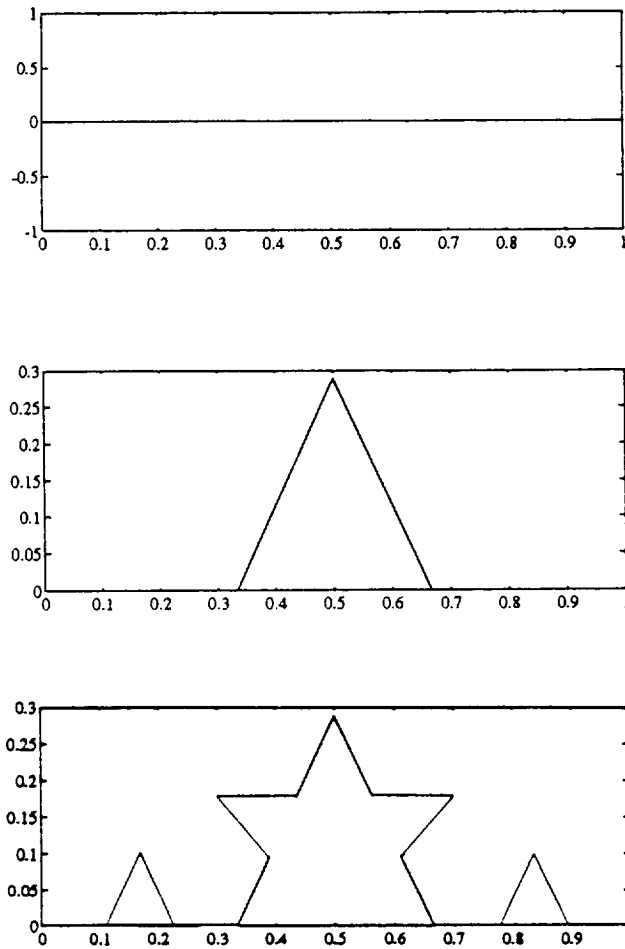


Figure 2.1 Koch curve as an example of a curve with fractional dimension. The first three steps in the construction of the Koch curve are shown. This process is repeated *ad infinitum* to generate the ideal Koch curve. The Hausdorff dimension for this curve is  $\frac{\ln 4}{\ln 3} = 1.2619$ .

points as the first curve. This is repeated until, at the  $n^{\text{th}}$  step, each linear segment of the curve is replaced by four segments, each of length  $\left(\frac{1}{3}\right)^n$ .

While this curve is obviously not one dimensional, neither does it appear two dimensional. In this case, one must resort to a more formal definition of dimension. The Hausdorff dimension of a general curve is figured in the following way: If  $D$  is the dimension of an object, then each time the scale of measurement is reduced by a factor  $a$ , coverage of the object requires  $a^D$  times as many unit measurement objects. Since at each step in the construction of the Koch curve the scale is reduced by a factor of  $a = 3$  and the number of objects for coverage increases by 4, the Hausdorff dimension is  $D = \frac{\ln 4^n}{\ln 3^n} = 1.2619$ . Objects with noninteger Hausdorff dimension have come to be generically termed “fractals” [26, 27].

The Hausdorff dimension can be used as a geometrical measure of how “strange” an attractor is, and has been used by various researchers to describe attractors. However, more recently some researchers have noted that a strictly geometrical measure of the attractor dimension may, on the one hand, not be the most appropriate measure and, on the other hand, is generally very difficult to calculate in practice [23, 4].

It has been suggested that a more appropriate measure of strangeness of attractors is one which takes into account not only the geometrical shape of the attractor, but also the frequency of visitation of the state to various portions of the attractor [4, 6]. One such measure is the *correlation dimension*, to be discussed in the following chapter.

### 2.2.2 Examples of Chaotic Systems

Several examples of chaotic systems have been discussed and explored in detail, both empirically and theoretically. These provide important frames of reference for discussion since they are so well understood.

A very trivial example of a chaotic system is the *discrete time logistic equation*, given by the following formula:

$$x[k + 1] = \mu x[k](1 - x[k]). \quad (2.1)$$

For parameter values of  $0 < \mu \leq 4$  and initial condition  $0 < x[0] < 1$  this system maps the unit interval into itself; for particular values of  $\mu$ , the sequence  $x[k]$  is chaotic. Figure 2.2 shows a typical sequence with  $\mu = 3.5699456$  (just at the onset of chaos) [5] and initial condition  $x[0] = 0.6791$ . The Hausdorff dimension of the attractor in this case is known to be 0.538 [5].

A second example, this one presented by Lorenz [1] for modelling convective fluid flow, is given in continuous time by the set of nonlinear, coupled differential equations

$$\begin{aligned}\dot{x} &= \sigma(y - x) \\ \dot{y} &= -xz + rx - y \\ \dot{z} &= xy - bz,\end{aligned}\tag{2.2}$$

where  $\sigma$ ,  $r$ , and  $b$  are constants and the dot denotes differentiation with respect to time. For particular values of these parameters the system exhibits chaotic behavior. Figure 2.3 shows a three dimensional plot of the attractor of the Lorenz system for values of  $\sigma = 10.0$ ,  $r = 28.0$ , and  $b = \frac{8}{3}$ , a set of parameter values which is known to result in chaotic behavior [5]. (The simulation was performed using the fourth order Runge-Kutta numerical integration method with time step 0.002. The first 2000 points were discarded to ensure that the trajectory had settled onto the attractor.) Figure 2.3 shows the three dimensional picture of 2000 points of a typical Lorenz attractor. This attractor is known to have Hausdorff dimension  $2.06 \pm 0.01$  [5].

A third system, known as the Rössler system, is given by the following set of differential equations:

$$\begin{aligned}\dot{x} &= -z - y, \\ \dot{y} &= x + ay, \\ \dot{z} &= b - cz + zx,\end{aligned}\tag{2.3}$$

which exhibits chaotic behavior for parameter values  $a = 0.15$ ,  $b = 0.20$ , and  $c = 10.0$ . An attractor for this system is shown in Figure 2.4.

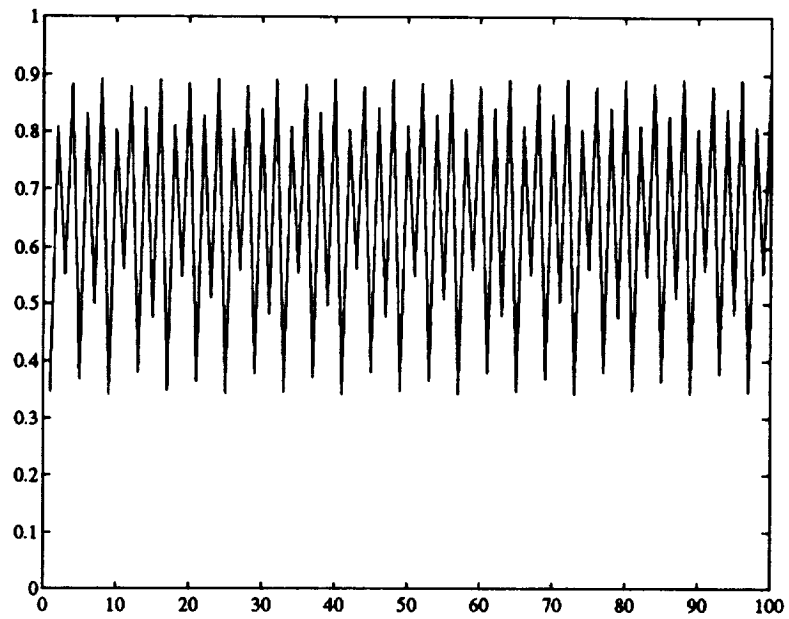


Figure 2.2 Sample attractor for the logistic equation with  $\mu = 3.5699456$  and initial condition  $x[0] = 0.6791$ .

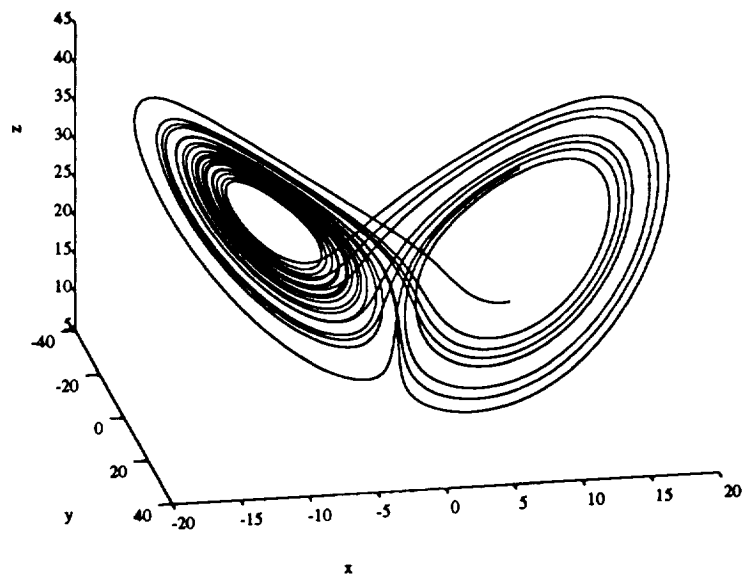


Figure 2.3 Sample Lorenz attractor simulated using fourth order Runge–Kutta numerical integration scheme. Parameter values are  $\sigma = 10.0$ ,  $r = 28.0$ , and  $b = \frac{8}{3}$ .

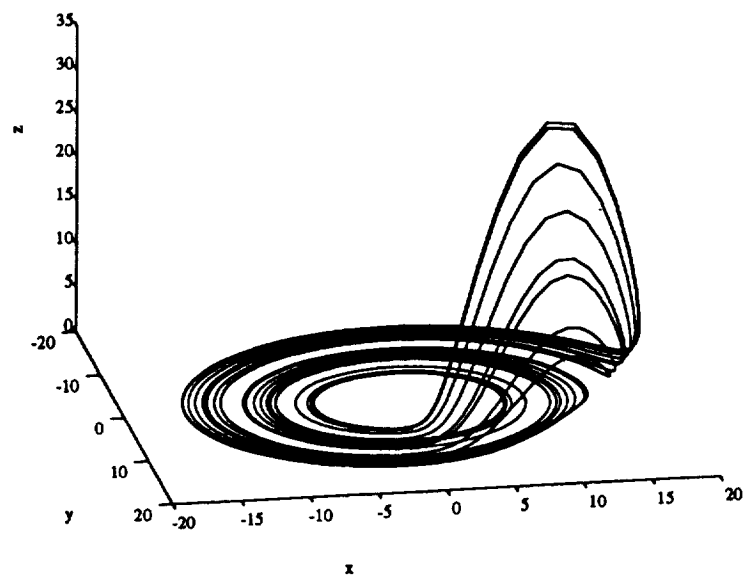


Figure 2.4 Sample Rössler attractor with parameter values  $a = 0.15$ ,  $b = 0.20$ , and  $c = 10.0$ .



## CHAPTER 3

### THE CORRELATION DIMENSION

This chapter begins a study of the characterization of an attractor based on its dimension, the types of dimension that may be used, and the physical significance of these dimensions. Geometrical, mathematical, and dynamical concepts will be used as aids in the development as necessary. Attention is focussed on the correlation dimension, since this dimension has been found to be simpler to calculate from a time sequence and more physically meaningful in most cases.

#### 3.1 Measures of Dimension

The purpose of the work here is to eventually gain some knowledge about a system based on a time sequence which is output from the system. The question is: What shall we use as a descriptor of the system? The descriptor should be an invariant measure of the system, *i.e.*, a measure which is independent of the specific orbit of the trajectory (or, equivalently, independent of the specific initial condition) [25], or invariant with time [12]. One such invariant, the one which will be considered here, is the dimension of the system's attractor. However, there are several choices for the dimension to be calculated. Below several measures of dimension will be defined and problems associated with either the measure or with its practical calculation will be discussed. As a practical matter, one would like a measure of an attractor which is both reasonable in computational complexity and as descriptive of both the attractor structure and the dynamics on the attractor as possible.

The idea that the Hausdorff dimension,  $D$ , of a chaotic attractor is noninteger was discussed in the previous chapter. In the past, this has been a popular feature used to characterize systems, but recently it has been realized that there are some flaws in its use. First, as noted in Chapter 2, it is a strictly geometric measure, simply finding

the spatial dimension occupied by the attractor. This gives no emphasis to portions of the attractor which are visited more often [23, 4]. From a dynamical systems point of view, this is important, since information is conveyed by the probability of finding the state in a given region, as well as by the geometric structure of the attractor.

Second, the Hausdorff dimension has been found to very difficult to calculate whenever the dimension is greater than 2. Further, the most popular method of finding the Hausdorff dimension is using box counting. This method has been found to be both very computational, since all calculations must be repeated every time the scale is changed, and very slow to converge, often requiring on the order of  $10^5$  points [23, 4]. The convergence problem is perhaps the most critical, since the number of points available is typically at least an order of magnitude less.

A second dimension is the information dimension, though there is no single accepted definition of this measure. The concept of information dimension results from viewing the system as a generator of information [6]. The exponential divergence of trajectories (determined by the positive Lyapunov exponents) means that information is created as each succeeding point on the attractor is known, since specification of each successive point more accurately specifies the initial condition. According to Farmer *et al.* [6], “The information dimension is a generalization of the capacity that takes into account the relative probability of the cubes used to cover the set.” Generally, the information dimension is taken as meaning something like the number of bits needed to specify a point on the attractor to a given accuracy [6]. (Here in particular there is no standard definition; Eckmann and Ruelle [23] refer to the correlation dimension which follows as “information dimension”.)

Another measure has been introduced which not only is easier to calculate, but also approximates (more precisely, lower bounds, but the bounds are usually very tight [5]) both the Hausdorff and information dimensions [5]. Further, this measure has been argued to be more physically meaningful than the Hausdorff dimension itself [4]. This measure is the *correlation dimension*, as discussed in the following section.

## 3.2 The Correlation Integral and Correlation Dimension

### 3.2.1 Definition of Correlation Integral

The correlation dimension is defined as a parameter of the correlation integral. The *correlation integral* is simply a type of empirical distribution function based on pairwise distances between pairs of vectors, calculated as follows: Given a time sequence of vectors  $\{\mathbf{x}[k]\}_{k=0}^{N-1}$ , the distances between all pairs of vectors are calculated (using either Euclidean or maximum distance measures). The correlation integral itself,  $C(r)$ , is calculated as a function of distance  $r$  as

$$C(r) = \lim_{N \rightarrow \infty} \frac{2}{N(N-1)} \sum_{i=1}^N \sum_{j=1}^{i-1} \Theta(r - |\mathbf{x}_i - \mathbf{x}_j|), i \neq j, \quad (3.1)$$

where  $\Theta(\cdot)$  is the Heaviside step function,

$$\Theta(a) = \begin{cases} 0, & a < 0 \\ 1, & a > 0 \end{cases}.$$

### 3.2.2 Shape of Correlation Integral

This shows that  $C(r)$  simply represents the fraction of vectors which lie within a distance  $r$  of one another. A few simple examples are considered to demonstrate the physical meaning of the correlation integral: first, a line containing points equally spaced from 1 to 1000. If this is an attractor embedded in, for example, three dimensional space, then the attractor has diameter  $999\sqrt{3}$  (using the Euclidean distance measure). Up to that diameter, the correlation integral scales linearly with  $r$ , *i.e.*,  $C(r) \propto r$ . Figure 3.1 shows the correlation integral, while Figure 3.2 shows the correlation dimension, which is indeed calculated to be nearly one. Similarly, if the attractor fills two dimensional space and the points are equally spaced, it would be found that, up to the diameter of the attractor,  $C(r) \propto r^2$ , giving a correlation dimension of 2. Note in these examples that the correlation dimension is the exponent of  $\nu$ . Further, note that while *any* line would have a Hausdorff dimension of 1, the points must be equally spaced on the line for the correlation dimension to also be 1,

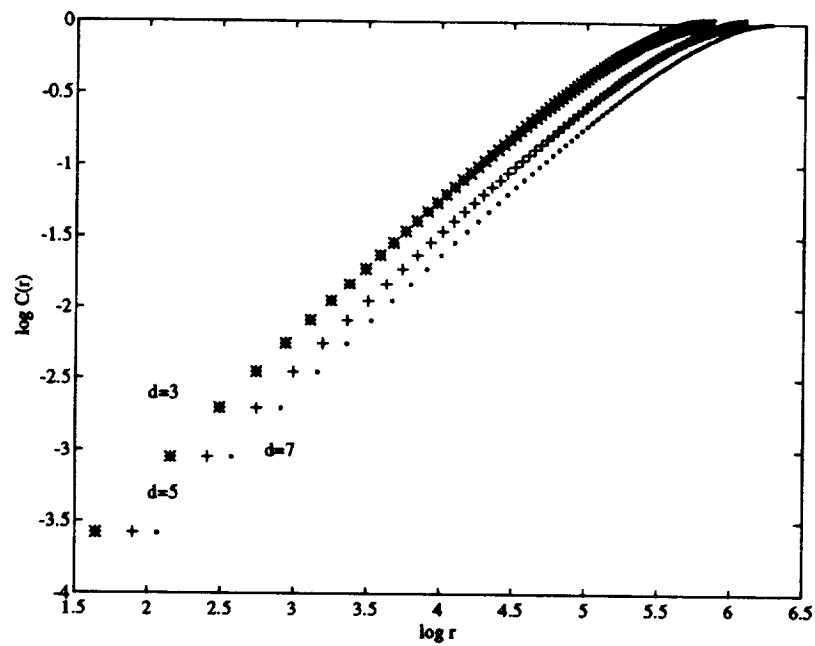


Figure 3.1 Correlation integral for linear attractor with equally spaced points.  
 $T = 7$  and embedding dimensions are 3, 5, and 7.

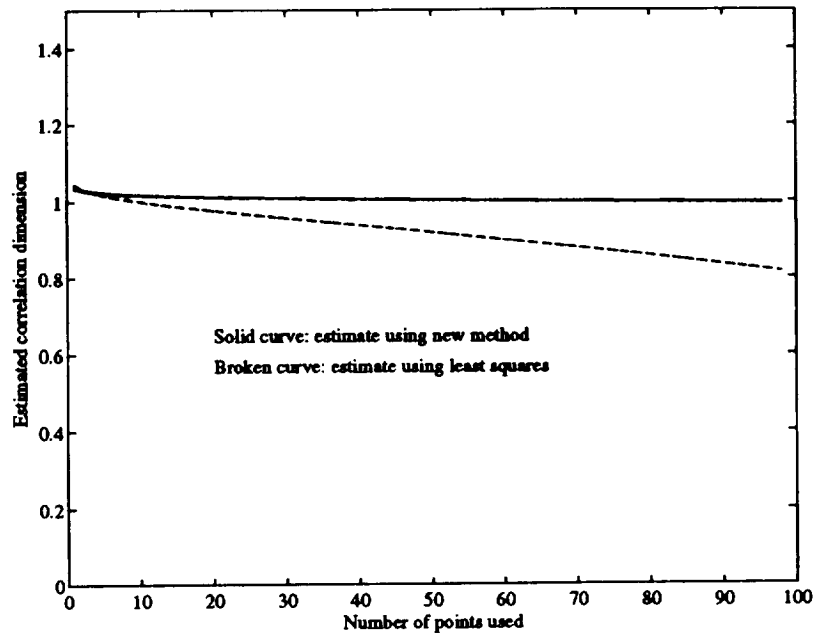


Figure 3.2 Estimates of the correlation dimension for linear attractor with equally spaced points.  $T = 7$  and embedding dimensions are 3, 5, and 7. The attractor is shown to have dimension one. See Chapter 4 for details on methods used to calculate the correlation dimension from the correlation integral.

since the correlation dimension compensates for the relative frequency of occurrence of points on the line.

This is true for attractors with fractal structure also; for small values of  $r$  the correlation integral has a power law dependence on the range  $r$ :

$$C(r) \propto r^\nu, \quad (3.2)$$

but the exponent  $\nu$  may be noninteger. The parameter  $\nu$ , referred to as the *correlation dimension*, can be found as the slope of the plot of  $\log[C(r)]$  versus  $\log[r]$  for small values of  $r$ . Mathematically,

$$\nu = \lim_{r \rightarrow 0} \frac{\log[C(r)]}{\log[r]}. \quad (3.3)$$

### 3.3 Reconstruction of System Dynamics from a Time Series

The overall goal is to derive some estimate of the dimension of the attractor for a chaotic system. In general, all that is available is a single time sequence, which is the result of measurement of some component or components of the system, and the system must be analyzed based solely on this sequence. However, if techniques such as correlation dimension analysis are to be applied to the system, a sequence of *state vectors* is required. As stated by Packard *et al.* [3], in geometric terms the problem is to reconstruct phase space from the observation of a single coordinate. Essentially, one is trying to guess both the dimensionality  $n$  of the attractor and the topology of the phase space [3]. In this section one method of deriving the required  $n$  independent quantities from a time series is explored.

#### 3.3.1 The Number of Independent Measurements Required

For a system of topological dimension  $n$ , it is well known that unique specification of the state of the system requires at least  $n$  *independent quantities*. This is logical, since the topological dimension is the topological dimension of the phase space, and the number of components required to uniquely specify a point in space is equal to the topological dimension of the space [3]. The *state* of the system is just a particular point

in phase space at a given time. While a system may be described in many ways in terms of the physical meaning of the components of the state vector (any state vector being diffeomorphically equivalent to any other), the common feature is that exactly  $n$  components are needed for any state vector description. (More components simply add redundant information; fewer components do not give a *unique* description.)

The key to the use of a single time series for reconstruction of dynamics is that *any  $n$  independent* quantities can be used as a state vector for the system. Much work has been concentrated on this problem in the last decade, with the basic idea that the measurements composing the time series are *independent* of one another (in a sense to be defined later) if measurements are taken which are spaced an appropriate distance apart (with *appropriate distance* also to be defined later). This general method is called the method of *time delay coordinates*, introduced by Packard *et al.* [3].

### 3.3.2 The Method of Time Delay Coordinates

While it may be unreasonable to expect to construct a state vector whose components have a physical significance directly, it has been argued that *any  $n$  independent* quantities from a system are sufficient to completely describe the system at any particular time. A choice for these quantities which has become popular with researchers, particularly for its simplicity, is time delay coordinates.

The idea behind this particular method is that since the system is dynamic, the measurements in the time sequence are changing. The vector

$$\mathbf{x}[k] = \begin{bmatrix} x[k] \\ x[k - T] \\ \vdots \\ x[k - (d - 1)T] \end{bmatrix} \quad (3.4)$$

is formed, where the boldface  $\mathbf{x}$  denotes a vector, the plain  $x$  denotes the scalar time sequence,  $T$  is an integer delay, and  $d$  is an integer called the *embedding dimension*. (Methods of choosing  $T$  and  $d$  will be discussed below. It is also possible to use the same method with samples taken at times  $k, k + T, \dots, k + (d - 1)T$ . This is simply

a matter of notation.) With appropriate values of  $T$  and  $d$ , the elements of  $\mathbf{x}$  will be independent and will contain enough information to specify the state of the system which generated the time sequence.

### 3.4 Relationship between Time Delay and Phase Space Reconstructions

#### 3.4.1 Influence of Time Delay on Coordinates' Independence

One might surmise that the choice of  $T$  in Equation 3.4 is critical in the success of any analysis based on time delay coordinates. Empirical evidence has shown that the method of time delay embedding is relatively robust with respect to the choice of  $T$ , within the limitations discussed below. In fact, even the fundamental work of Fraser and Swinney on choosing  $T$  states that “For an infinite amount of noise free data, the time delay  $T$  can in principle be chosen almost arbitrarily” [9]. However, that scenario cannot be assumed in the real world, so this will be explored further. Basic to the choice of value of  $T$  is that the components of  $\mathbf{x}$  be independent but not unrelated. Two questions should be posed: what are the implications if  $T$  is not chosen properly, and what definition of *independence* should be used?

It is helpful to take a geometrical view in order to answer the first question. Successive components of the time sequence are generally highly correlated in a statistical sense; if there is no correlation, the sequence is white noise by definition. Correlation between adjacent samples also implies correlation between samples spaced farther apart in time; in empirical studies, determining independence involves testing when the (auto)correlation drops below and stays below a certain threshold. The implication is that if  $T$  is chosen too small, say  $T = 1$ , the components of  $\mathbf{x}$  are highly correlated, so that the attractor seems to be a line in phase space. Figure 3.3 shows the partial phase space reconstruction for the Lorenz system obtained when the time delay is too small.

As a guideline for choice of the upper bound on  $T$ , it should be recalled that at least one Lyapunov exponent of a chaotic system is positive. This means that



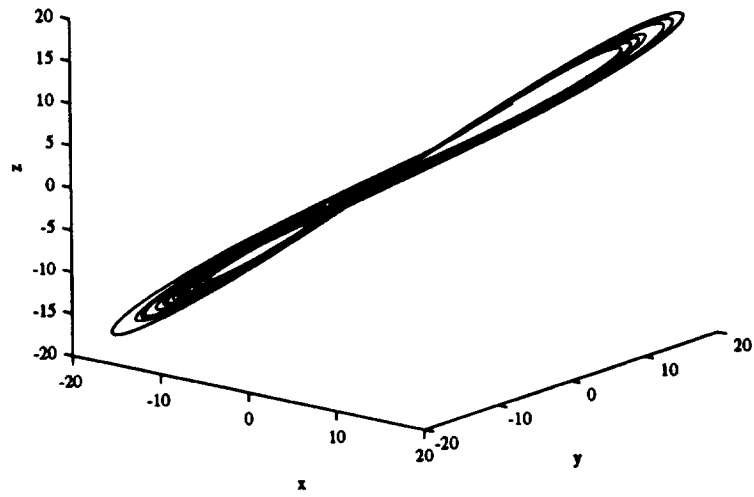


Figure 3.3 Reconstruction of Lorenz attractor using time delay coordinates with time delay too small,  $T = 1$  in this case. This causes the attractor to appear to lie along the line  $x = y = z$ .

the trajectory in phase space is exponentially diverging in at least one direction. Thus, if  $T$  is chosen too large, the components of  $\mathbf{x}$  will be unrelated due to the dynamical nature of the system and the divergence of trajectories. Figure 3.4 shows the consequences of choosing  $T$  too large; the reconstructed attractor does not show the rich structure of a typical Lorenz attractor, but looks rather jumbled.

This discussion is useful in understanding the independence concepts involved, but researchers have recently noted [9, 28] that this definition of independence based on a threshold on the autocorrelation only implies *linear* independence. This is not appropriate in the study of chaotic systems, since these systems are inherently *non-linear*. The proposal has been to use *general* independence [9], which is measured in an information theoretic sense [29]. Mathematically, rather than looking at the autocorrelation function, we look for the value of  $T$  which yields the *first minimum* of the *mutual information function* [29, 9]

$$I(T) = \sum_k P\{x[k], x[k-T]\} \log \left( \frac{P\{x[k], x[k-T]\}}{P\{x[k]\}P\{x[k-T]\}} \right). \quad (3.5)$$

Arbanel [25] points out that choosing the first minimum of this function with respect to  $T$  is the “...natural, nonlinear, information theoretic analog of choosing the first zero of the autocorrelation function as a correlation time and useful time delay.”

The mutual information between delayed versions of the  $x$  component of the Lorenz system is shown in Figure 3.5. Noting that the first minimum occurs at  $T = 17$ , this is chosen as the appropriate delay for the time delay embedding. Figure 3.6 shows the phase space reconstruction using the sequences  $x[k]$ ,  $x[k - 17]$ , and  $x[k - 2 * 17]$  for the  $x$ ,  $y$ , and  $z$  components, respectively. What is important here is not that this phase space reconstruction exactly resemble the original Lorenz attractor (such as the example shown in Figure 2.3), but that the components be “independent” yet not unrelated.

Equation 3.4 shows that the time delay vector  $\mathbf{x}$  has  $d$  components. Since the state vector should *uniquely* specify the state of the system, it must contain at least as

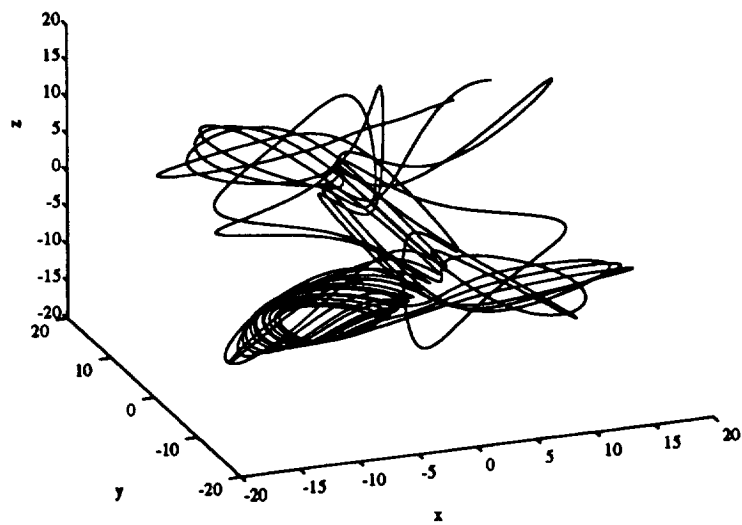


Figure 3.4 Reconstruction of Lorenz attractor using time delay coordinates with time delay too large,  $T = 77$  in this case. This causes the coordinates to be unrelated and the attractor to appear random.

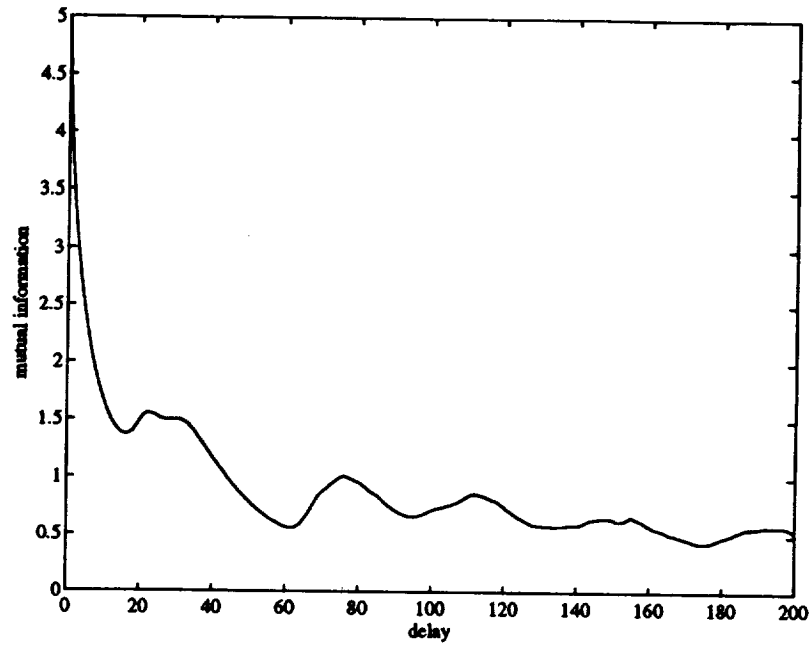


Figure 3.5 Mutual information between delayed versions of the x component of the Lorenz system. The first minimum occurs at  $T = 17$ , so this is accepted as the appropriate delay for the time delay embedding.

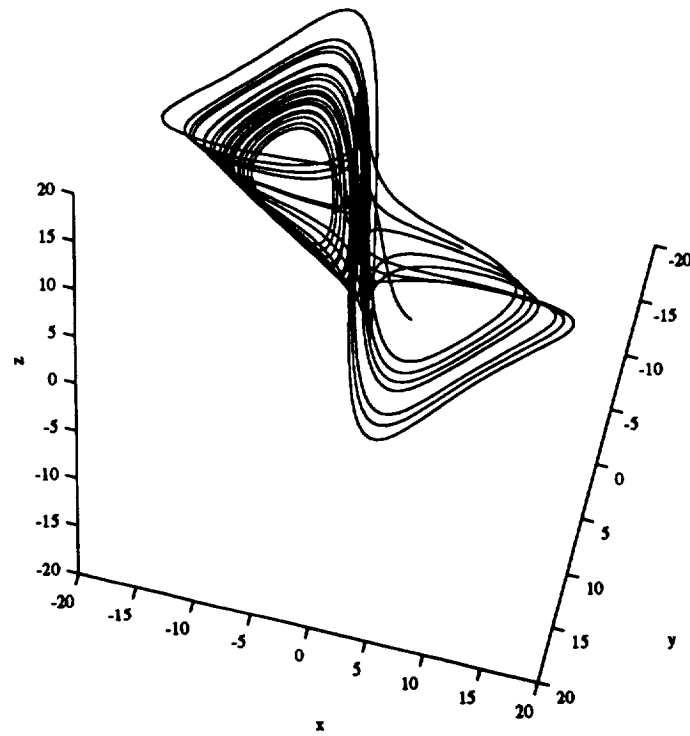


Figure 3.6 Reconstruction of Lorenz attractor using time delay coordinates with time delay chosen as the first minimum of the mutual information function,  $T = 17$ .

many components as there are degrees of freedom in the system. Therefore, naively we can say that  $d$  must be at least as large as the topological dimension of the system's attractor (which is the same as the number of degrees of freedom of the system). The Takens Embedding Theorem [7] proves that phase space can be effectively reconstructed based on vectors of dimension  $d \geq 2d_{\text{actual}} + 1$ . Arbanel [25] says that this bound on  $d$  is sufficient but not necessary. He (and others) uses the idea of orbit (trajectory) crossings in the determination of  $d$ .

What is being described is essentially the reconstruction of *a* phase space using vectors derived from one time sequence. (Emphasis is on “a” phase space, since the state vector description of a system is not unique. As stated earlier, the state vector descriptions are diffeomorphically equivalent, however.) It is appropriate to note here that not all reconstructions of phase space are equally desirable. Fraser [28] comments that “bad” reconstructions are those which are not invertible, or in which “... points in the reconstructed phase space do not uniquely specify points in the original phase space.” However, this is more a theoretical than a practical matter, and in practice a choice of  $T$  which gives *any* reasonable reconstruction is sufficient.

### 3.4.2 A Conjecture on the Appropriateness of Uniform Time Delay Based on Mutual Information

The work of Fraser and Swinney [9] was the first and most significant attempt to put the choice of time delay for embedding on a firm theoretical footing. They conclude that the delay which gives the first minimum of the mutual information function is the appropriate delay for time delay embedding. However, while there seems great theoretical justification for this choice, a couple of problems will be noted and questions posed.

While this choice indeed leads to the first two coordinates (and indeed all pairs of successive coordinates) being generally independent, it ignores significant dependence between other pairs of coordinates. Figure 3.7 shows the mutual information between delayed versions of the  $x$  component of the Rössler system of Equation 2.3. At a

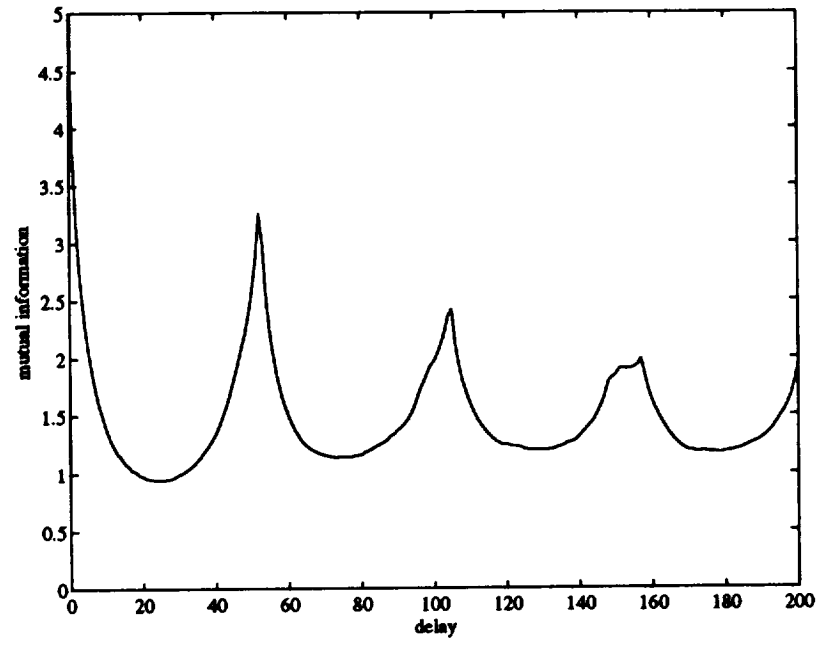


Figure 3.7 Mutual information between delayed versions of the x component of the Rössler system. First minimum is at  $T = 25$ .

delay of  $T = 25$  the mutual information between contiguous pairs of coordinates in the time delay vector will be very small; however, at a delay of  $2T = 50$ , the mutual information will be much larger. This means that components  $x_1, x_3, x_5$ , etc. will be very dependent, as will be components  $x_2, x_4$ , etc. It seems more appropriate to consider general independence between *all* the coordinates together. In this sense, a more sensible criterion than the mutual information of Equation 3.5 is the general relation

$$I(T) = \sum_k P\{x[k], x[k-T], \dots, x[k-(d-1)T]\} \times \\ \log \left( \frac{P\{x[k], x[k-T], \dots, x[k-(d-1)T]\}}{P\{x[k]\}P\{x[k-T]\} \dots P\{x[k-(d-1)T]\}} \right).$$

Implementation of such a criterion would be very impractical, since the co-occurrence probabilities of large combinations of values would be required, particularly when the embedding dimension  $d$  is large. This poses an extreme computational burden, as well as statistical problems, for the co-occurrences will be very sparse. Thus, as a practical matter, the computation would be prohibitive and this will not be pursued further here.



## CHAPTER 4

### A NEW ESTIMATE FOR THE CORRELATION DIMENSION

#### 4.1 Traditional Method Using Least Squares

The method of least squares is a common tool for fitting a curve through a finite and known number of data points. (In the special case that the curve is linear, as here, the term "linear regression" is often used.) The method is well known and understood, but the development will be given here for comparison with the new method that follows.

A set of data is available with independent variable  $x_i$  and dependent variable  $y_i$ , with  $1 \leq i \leq N$ , so that there are a total of  $N$  pairs of coordinates  $(x_i, y_i)$ . For the linear regression case, the assumed form of the curve is  $y = ax + b$ , so that  $a$  and  $b$  are parameters which are determined as follows. The error criterion chosen is that the sum of squared errors between the curve and the data points is minimum, or that

$$e^2 = \sum_{i=1}^N [y_i - (ax_i + b)]^2 \quad (4.1)$$

is minimized by the choice of  $a$  and  $b$ .

Minimization of Equation 4.1 is accomplished in the standard way by taking partial derivatives with respect to  $a$  and  $b$  and setting to zero, which yields

$$\frac{\partial e^2}{\partial a} = \sum_{i=1}^N 2[y_i - ax_i - b](-x_i) = 0$$

and

$$\frac{\partial e^2}{\partial b} = \sum_{i=1}^N 2[y_i - ax_i - b](-1) = 0.$$

These reduce to

$$a = \frac{\sum_{i=1}^N x_i y_i - b \sum_{i=1}^N x_i}{\sum_{i=1}^N x_i^2} \quad (4.2)$$

and

$$b = \frac{1}{N} \left[ \sum_{i=1}^N y_i - a \sum_{i=1}^N x_i \right]. \quad (4.3)$$

Combining equations 4.2 and 4.3 gives the final result:

$$a = \frac{\sum_{i=1}^N x_i \sum_{i=1}^N y_i - N \sum_{i=1}^N x_i y_i}{\left[\sum_{i=1}^N x_i\right]^2 - N \sum_{i=1}^N x_i^2}. \quad (4.4)$$

(Note that the application here only involves the slope, so this will not be continued further to find the intercept.)

## 4.2 A New Method

It has been noted by other researchers that individual points of the correlation integral are not independent in a statistical sense. This is because of the nature of the correlation integral: it is essentially an empirical distribution function. As noted by Theiler [10], the value of  $C(r + \varepsilon)$  is just  $C(r)$  plus the fraction of distances between  $r$  and  $r + \varepsilon$ . Thus the values are highly dependent. Least squares methods are not optimal for use when the data points are not independent.

### 4.2.1 Error Criterion and Justification

On the other hand, the individual *increments* of the correlation integral are independent. Therefore, here it is argued that a more sensible error criterion might be

$$e^2 = \sum_{i=2}^N \{y_i - y_{i-1} - [(ax_i + b) - (ax_{i-1} + b)]\}^2$$

or

$$e^2 = \sum_{i=2}^N [y_i - y_{i-1} - a(x_i - x_{i-1})]^2. \quad (4.5)$$

Taking the derivative of Equation 4.5 with respect to  $a$  and setting to zero gives

$$a = \frac{\sum_{i=2}^N (x_i - x_{i-1})(y_i - y_{i-1})}{\sum_{i=2}^N (x_i - x_{i-1})^2}. \quad (4.6)$$

In this application, let  $x = \log(r)$  and  $y = \log[C(r)]$ , so that the slope  $a$  is the correlation dimension  $\nu$ .

#### 4.2.2 Advantages over Other Methods

The method for finding the correlation dimension discussed above has significant advantages over previous methods. Unweighted least squares is not justified theoretically for use with data which are highly correlated. Weighted least squares can compensate for this, but it has been found empirically that results from weighted least squares are little if any better than those using unweighted least squares [11]. The optimal estimator developed by Takens, given by

$$\hat{\nu} = \frac{-1}{\langle \log[\frac{r_{i,j}}{r_0}] \rangle},$$

while greatly justified theoretically [8], often gives results which are no better than those using unweighted least squares [11].

In addition, all methods require that one choose a region of the  $\log[C(r)]$  versus  $\log(r)$  curve over which the curve is acceptably linear. Least squares methods will underestimate the slope if one uses points beyond the “knee” of the curve, while statistical confidence is low if too few points are used. Also, noise effects at small  $r$  values will lead to inflated slope estimates when using only a few points in the calculation.

The calculation of  $\nu$  represented by Equation 4.6 seems more robust than these. First, since it is based on information in the *increments* of the correlation integral, there is no inherent statistical dependence between the data points. Second, the shape of the correlation integral seems particularly suited for this method. Since the correlation integral levels off as the range  $r$  approaches the diameter of the attractor, the increments in the dependent variable decrease, so that the contributions to the correlation dimension estimate in Equation 4.6 diminish to zero. This has several implications.

The inherent decrease in weighting at large ranges which the new method provides means that when no noise is present in the data, the correlation dimension estimate will saturate and there is no need to specify a stopping point on the correlation

integral. A more formal analysis follows: First, Equation 4.6 is repeated explicitly in terms of  $\nu$ , but retaining  $x = \log[r]$  and  $y = \log[C(r)]$ , with the increments  $x_i - x_{i-1}$  and  $y_i - y_{i-1}$  replaced by  $\Delta x_i$  and  $\Delta y_i$ , respectively:

$$\nu = \frac{\sum_{i=2}^N \Delta x_i \Delta y_i}{\sum_{i=2}^N [\Delta x_i]^2}. \quad (4.7)$$

Note that both  $\Delta x_i$  and  $\Delta y_i$  are nonnegative. Assume that the curve is linear up to  $N_1$ ,  $C(r) \propto r^\nu$  for  $1 \leq i \leq N_1$ , and that the curve stays below that line beyond that point, as shown in Figure 4.1. Then the estimate for the correlation dimension,  $\nu_{est}$ , is

$$\nu_{est} = \frac{\sum_{i=1}^N \Delta x_i \Delta y_i}{\sum_{i=1}^N [\Delta x_i]^2}. \quad (4.8)$$

Because  $C(r) \propto r^\nu$ ,  $1 \leq i \leq N_1$ ,  $\Delta y_i = \nu \Delta x_i$  over this region. With the form assumed above,  $\Delta y_i < \nu \Delta x_i$ ,  $N_1 < i \leq N$ . Then Equation 4.8 becomes

$$\nu_{est} \leq \frac{\nu \sum_{i=1}^N \Delta x_i \Delta x_i}{\sum_{i=1}^N [\Delta x_i]^2},$$

which implies  $\nu_{est} \leq \nu$ . Therefore, if the log-log plot of the correlation integral satisfies the form described above, this estimator will not overestimate the true correlation dimension.

The fact that the contributions diminish at larger ranges is beneficial in stabilizing the estimate of  $\nu$ , but there is also a drawback: noise. Since the effect of noise is to inflate the slope of the correlation integral at small ranges, the new method is not expected to perform well in the presence of noise. The effect of noise will be to cause the new method to seriously overestimate  $\nu$  at small ranges, and the estimate will probably not recover. More will be said about this in the following section, which is an empirical study of the preceding.

### 4.3 Comparison of Results: New Method Versus Least Squares

In the following, results of the correlation dimension estimation will be presented. Each figure contains two curves plotted versus the stopping point of the correlation

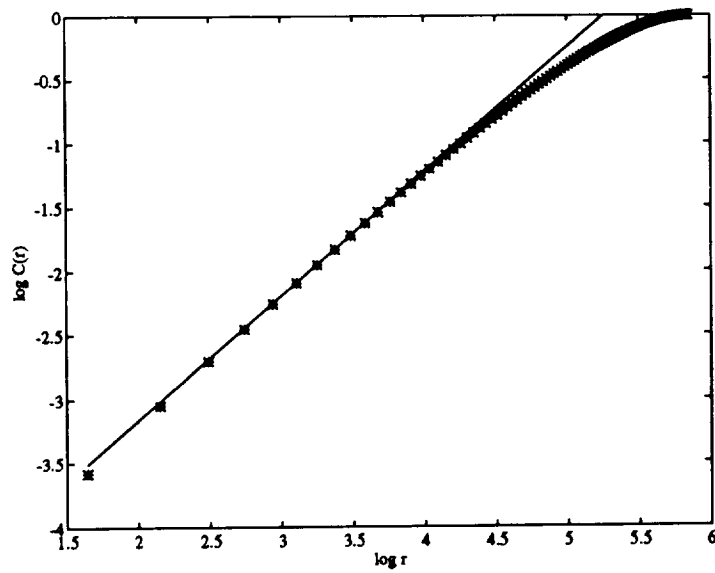


Figure 4.1 Assumed form for a typical  $\log[C(r)]$  *vs.*  $\log[r]$  curve. In this case, the estimator given for the correlation dimension will never overestimate the slope, regardless of what stopping point is chosen.

integral used for the estimate. One curve in each plot is derived using the least squares method and the other curve in each plot uses the new method. The results show that while the least squares estimate is strongly dependent on the stopping point chosen, the new estimate is practically unaffected by the choice.

Being based on the increments of the correlation integral, the new estimate for  $\nu$ , Equation 4.7, is more sensitive at small ranges, where the increments are larger. As the increments become smaller, the effect on the new estimate is smaller, whereas the least squares method has no inherent provision for such weighting.

It is known that the effect of noise in sequences is to inflate the slope of the correlation integral at small range values [11]. Thus, the new method is not expected to perform well in the presence of noise, since the noise will tend to dominate in the estimation of  $\nu$ .

Three studies are presented of simulated sequences whose correlation dimensions are known: the logistic equation (given by Equation 2.1 with  $\mu = 3.5699456$ , with a typical attractor as shown in Figure 2.2), the Lorenz system (Equations 2.2, with parameter values  $\sigma = 10.0$ ,  $r = 28.0$ , and  $b = \frac{8}{3}$ , and attractor shown in Figure 2.3), and the last involving a zero mean, unit variance Gaussian noise sequence. While the correlation dimensions of the first two examples are known, the correlation dimension of the white noise sequence is expected to follow the embedding dimension [10].

Figure 4.2 is a plot of the correlation integral for the logistic equation, while Figure 4.3 shows the estimated correlation dimension of the logistic equation versus the number of points of the correlation integral used for both the least squares method and the new method. The correct correlation dimension is known to be between 0.4926 and 0.5024 [5]. The least squares gives wildly fluctuating estimates, apparently due to the strange shape of the correlation integral, while estimates with the new method are very near the true value.

Figure 4.4 is a plot of the correlation integral for the Lorenz system of equations. Figure 4.5 shows the estimated correlation dimension of the system versus the number

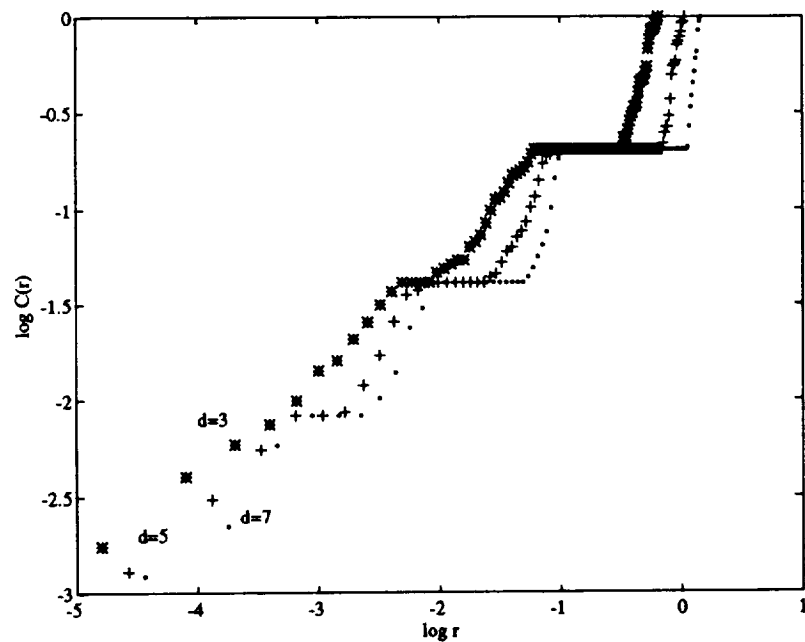


Figure 4.2 Correlation integral for the logistic equation. Time delay is  $T = 7$  and embedding dimensions are 3, 5, and 7.

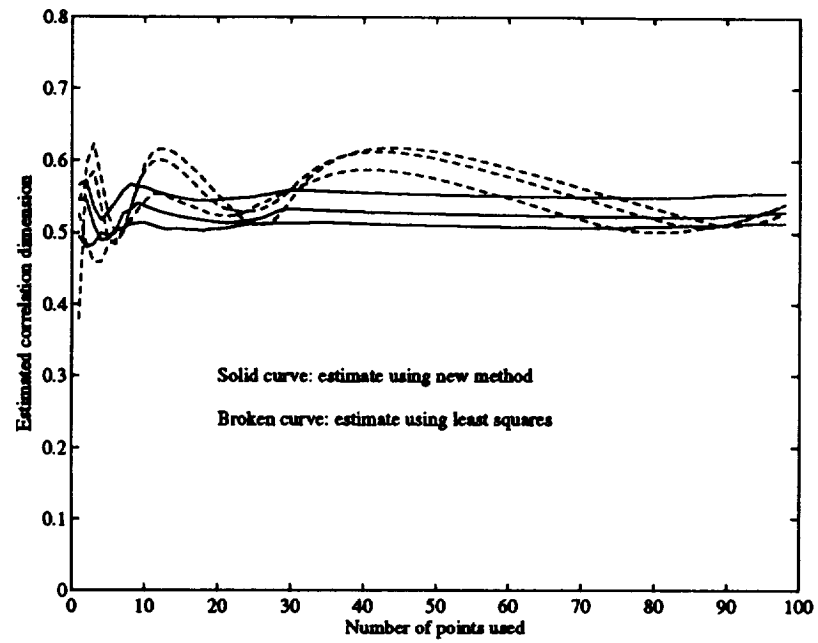


Figure 4.3 Estimates of the correlation dimension of the logistic equation using both the new method and least squares. True value is between 0.4926 and 0.5024.



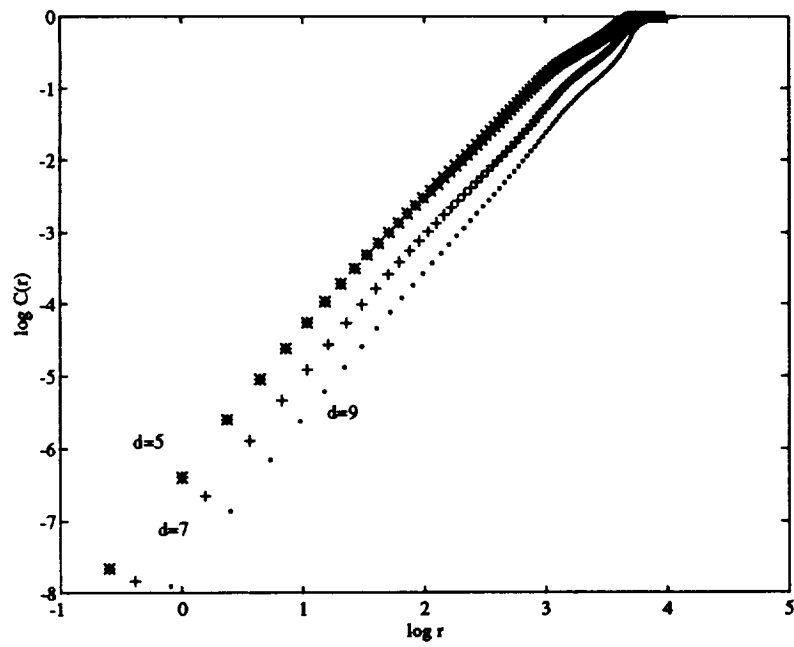


Figure 4.4 Correlation integral for the Lorenz system.  $T = 17$  and embedding dimensions are 5, 7, and 9.

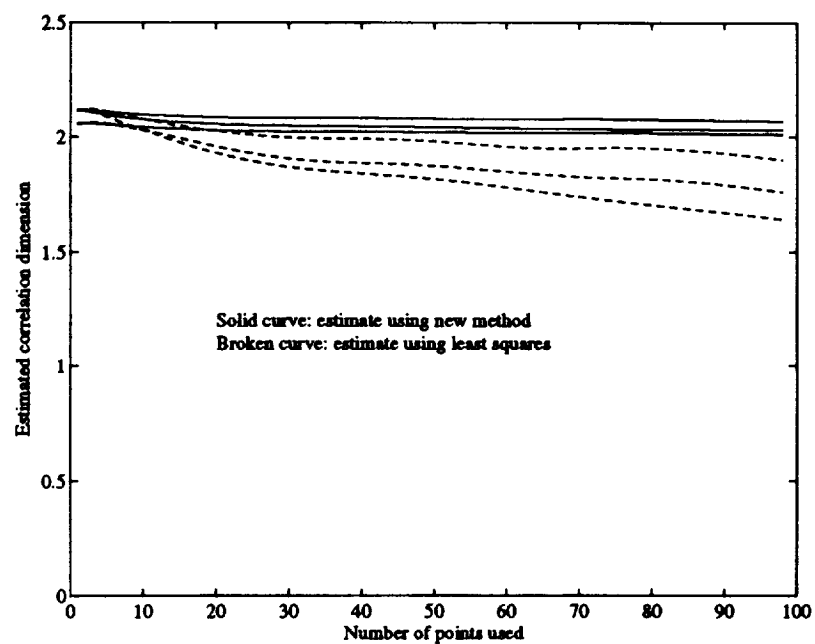


Figure 4.5 Estimates of the correlation dimension of the Lorenz system.  $T = 17$  and embedding dimensions are 5, 7, and 9. True value is approximately 2.05.

of points of the correlation integral used. The correct correlation dimension for the set of parameter values chosen is known to be  $2.05 \pm 0.01$  [5]. In this case, the least squares method gives an estimate which diminishes as more points are considered. (Note that this implies little statistical confidence can be put in the estimate, since it is only near the correct value when few points are used.) Again, the new method gives estimates which are very near the true value and are relatively unaffected by the stopping point.

Figure 4.6 shows the correlation integral for the Gaussian noise sequence for embedding dimensions of 3, 5, and 7. As expected, the slope of each increases with increasing embedding dimension. The calculated correlation dimensions for these curves are shown in Figure 4.7. In each case, the estimates for both the old and new methods are shown; the estimates begin near the same value, but the new method shows relative insensitivity to the stopping point used, while the least squares estimates decrease as the stopping point increases.

As noted earlier, the new estimator for the correlation dimension is not expected to perform well in the presence of noise. Figure 4.8 shows a plot of the correlation integral for the Lorenz system with zero mean, unit variance Gaussian noise added. The time delay used is  $T = 17$ , the correct delay based on the mutual information criterion. Embedding dimensions are 5, 7, and 9. Estimates of the correlation dimension from this correlation integral are shown in Figure 4.9 for both least squares and the new estimator. The randomness added to the attractor is expected to increase the correlation dimension estimates. Figure 4.9 shows that the noise effects on the new estimator are similar to its effects on the linear regression method.

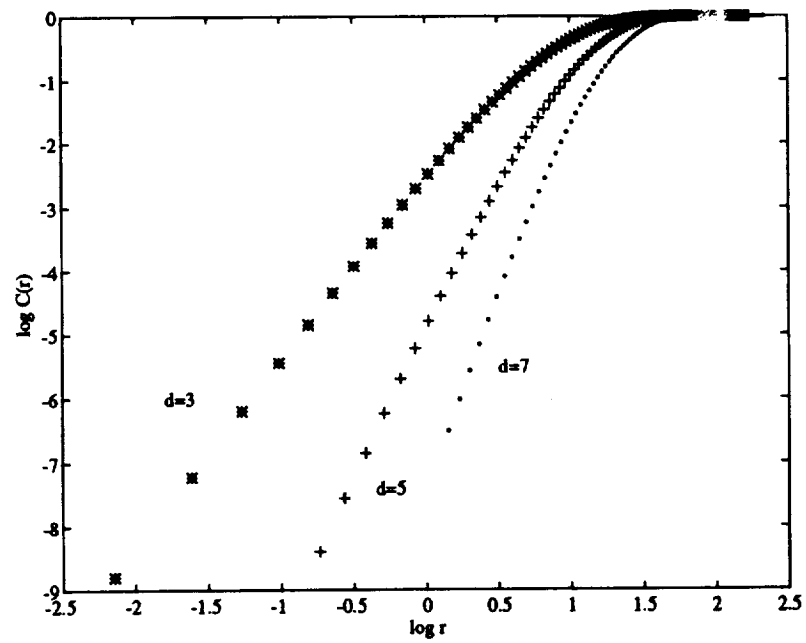


Figure 4.6 Correlation integral for the Gaussian noise sequence.  $T = 9$  and embedding dimensions are 3, 5, and 7.

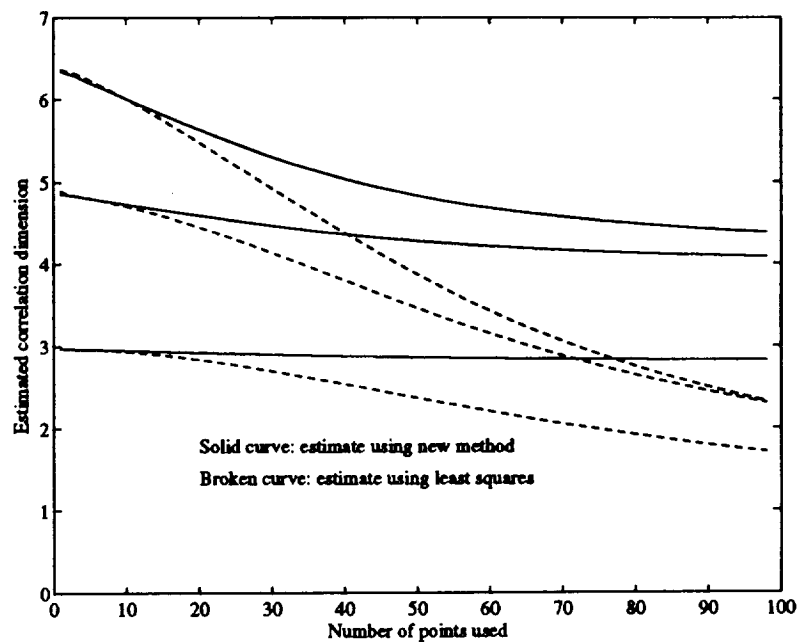


Figure 4.7 Estimates of the correlation dimension for the Gaussian noise sequence using both the new method and least squares.  $T = 9$  and embedding dimensions are 3, 5, and 7. Correlation dimension should equal embedding dimension in each case.

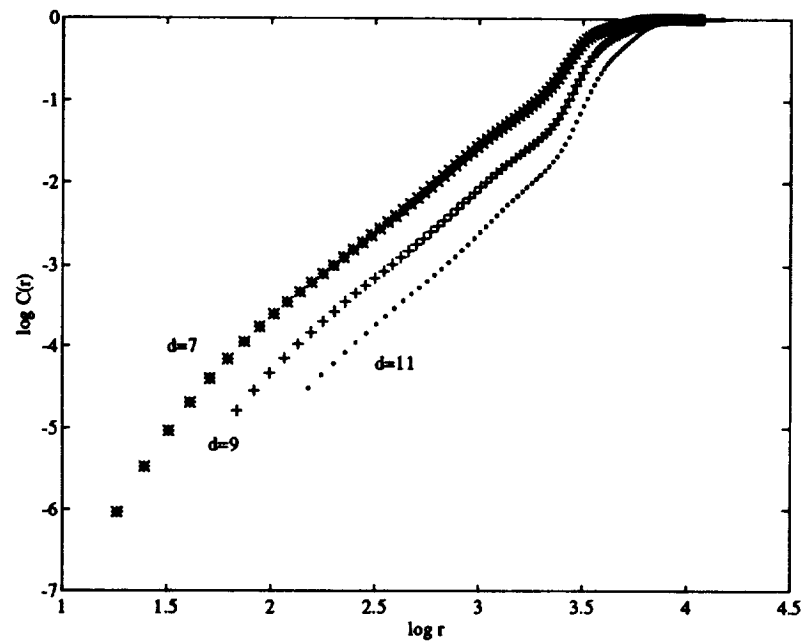


Figure 4.8 Correlation integral for the Lorenz system with zero mean, unit variance Gaussian noise added.  $T = 17$  and embedding dimensions are 7, 9, and 11.

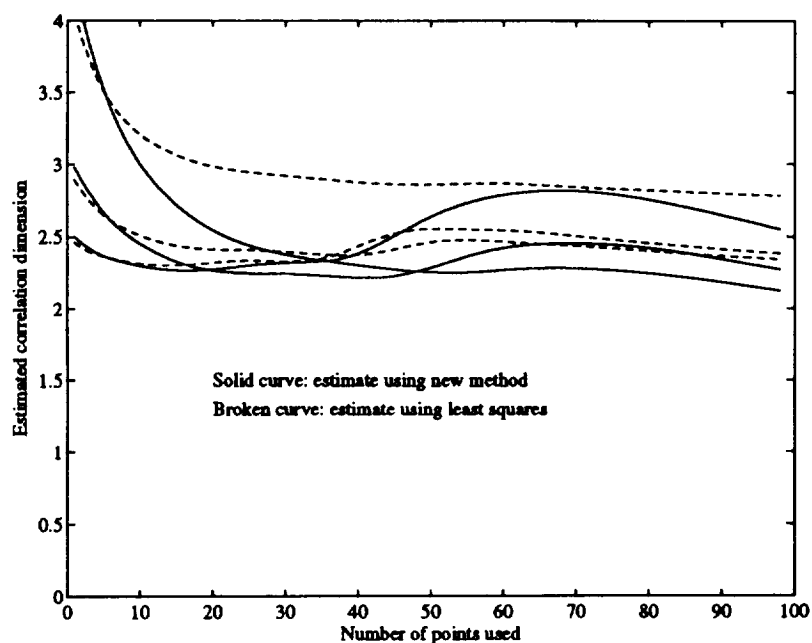


Figure 4.9 Estimates of the correlation dimension for the Lorenz system with zero mean, unit variance Gaussian noise added using both the new estimate and least squares.  $T = 17$  and embedding dimensions are 7, 9, and 11. Both methods are expected to overestimate the correlation dimension for the Lorenz system due to the presence of the noise.

## CHAPTER 5

### APPLICATION OF THE CORRELATION DIMENSION TO RADAR REFLECTOR DISCRIMINATION

One recent effort in the radar signal processing field is that of detecting low altitude windshear using airborne radar. This presents a multitude of problems, which arise primarily from the instabilities associated with radar platform motion and high return levels from ground clutter. Baxa [22] has identified the most significant problem as that of discriminating between ground clutter and slow moving weather, since frequency selective filters are ineffective in this case.

Lorenz showed that weather can be modelled as a chaotic system of relatively low dimension [1]. It is proposed here that this low dimension may lend aid in discriminating between radar returns from the weather and those from the ground. Previously, Leung and Haykin applied a similar idea to the determination of the correlation dimension of sea clutter [13, 17], concluding that the dimension of sea clutter is approximately 6.5. However, this was not compared to other types of reflectors. Two questions are to be answered: First, is the correlation dimension of weather returns truly “small” relative to that of other reflectors? And second, can the correlation dimension be used effectively as a discriminant between weather and ground returns?

#### 5.1 Radar System Scenario

The radar data were collected using an airborne radar system which scans in azimuth at  $0.5^\circ$  increments, collecting returns in 91 range cells at each azimuthal direction. Each return is a discrete time sequence of complex  $IQ$  data of length 96. The returns are converted to real sequences of length 192 samples by a complex modulation using the Discrete Fourier Transform. Since the  $3\text{ dB}$  beamwidth of the



antenna is approximately  $3^\circ$ , there is much spatial overlap in the field of view at different azimuths. This overlap gives justification for using averaging of the correlation dimension estimates over many range cells, since there is much redundancy in the physical space covered by the data.

Airborne radar data are available from two flights: The first is a clutter only flight over the Denver Stapleton Airport, while the second is a flight through a severe thunderstorm near the Orlando International Airport. The clutter is considered typical of ground clutter likely to be encountered near airports. Sidelobe returns are eliminated from the data by choosing range cells far enough ahead of the airplane that mainlobe returns dominate.

Two studies are presented here for both the weather and ground clutter returns: in the first, range cells 20 through 24 were chosen and followed over 200 azimuthal scans, giving a total of 1000 estimates of the correlation dimension. Afterward, some smoothing was performed on the estimates to reduce the effects of using such short time sequences in the correlation dimension estimation.

For the second study, range cells 30 through 36 were chosen and only the scans within  $\pm 1.5^\circ$  of the aircraft heading were used, giving a total of 7 scans per each complete scan in azimuth. These were followed over 21 total azimuthal scans, to give a total of 1000 estimates of the correlation dimension. The same type of smoothing was performed on this test.

## 5.2 Method of Analysis

The general technique for determining the correlation dimension of a system from a time sequence is as follows: Vectors are formed from the time sequence using time delay embedding, as described earlier. The correlation dimension is calculated from these vectors, increasing the embedding dimension  $d$  until there is saturation in the correlation dimension estimates. The value at which saturation occurs is then the correlation dimension of the system. Theiler has pointed out [11] that in such

discrimination problems, obtaining the *exact* correlation dimension is not as important as obtaining precise and repeatable estimates which allow discrimination.

### 5.3 Results of Correlation Dimension Analysis

Results of the correlation dimension calculations for the first case are shown in Figure 5.1 for the clutter only case and Figure 5.2 for the weather only case. Figures 5.1 and 5.2 show a marked difference in the correlation dimension of the weather and ground clutter returns. While saturation of the correlation dimension of the ground returns occurs around 3–4.5, the weather returns show saturation in the neighborhood of 2 – 3. The overall means for these estimates for the correlation dimensions are 2.3 for the weather and 3.5 for the ground return.

This difference is also evident in Figure 5.3, a clutter only scan, and Figure 5.4, which contains weather in the earlier scans but clear air (or ground) returns in the later scans. (Verification of the types of targets for the second test was done using measured reflectivities. The only high reflectivities encountered during the high altitude flight over Orlando were from weather.) Here the overall means for the estimates are 3.0 for the Orlando flight (the mean is shifted upward by the presence of nonweather return in the later scans) and 3.8 for the ground return.

Figures 5.5 and 5.6 show the probability of occurrences of the correlation dimension estimates for the two studies above. It can be seen that, particularly in the first case, the modes of the two distributions are well separated, so that discrimination based on a threshold on the correlation dimension would be very effective. In fact, if a threshold is set at 2.78, with correlation dimensions above the threshold considered ground and those below weather, a detection probability of 0.93 is possible while a false alarm rate of less than 0.10 is maintained.

For the second case, the threshold must be set at 3.70 to achieve a false alarm less than 0.10; with this threshold, the detection probability is 0.71. However, it must be kept in mind that the “weather” return here is also contaminated by clear air or

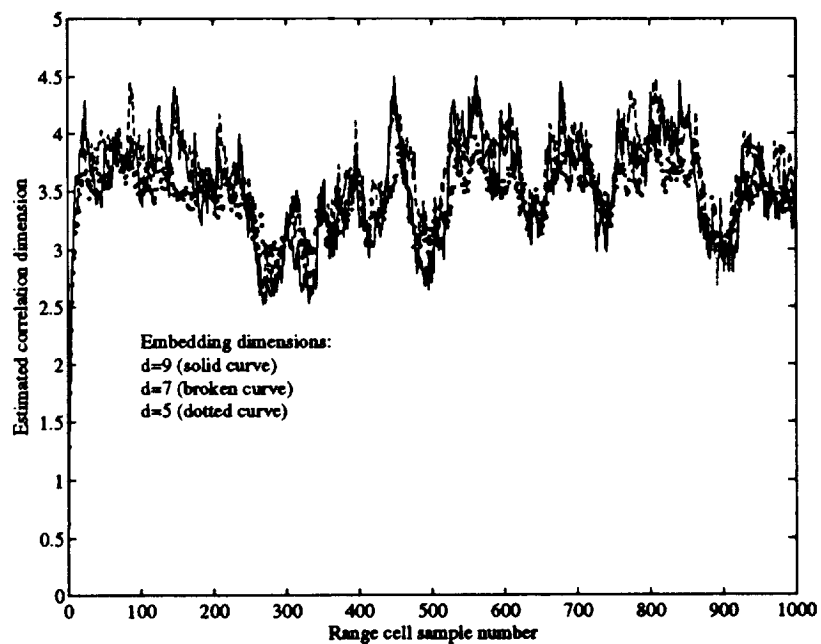


Figure 5.1 Calculated correlation dimension for ground clutter for the first test.  
Time delay is 5 samples. Embedding dimensions are 5, 7, and 9.

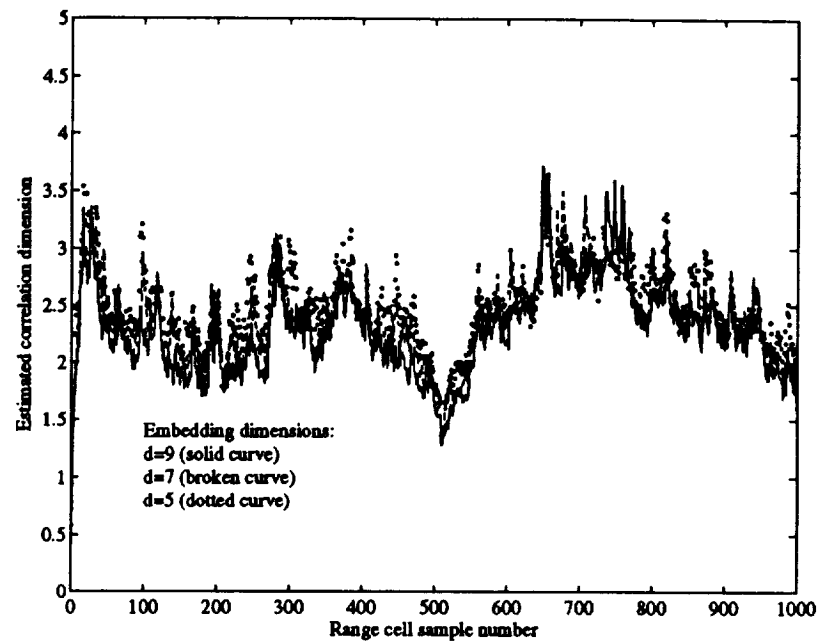


Figure 5.2 Calculated correlation dimension for weather for the first test. Time delay is 5 samples. Embedding dimensions are 5, 7, and 9.

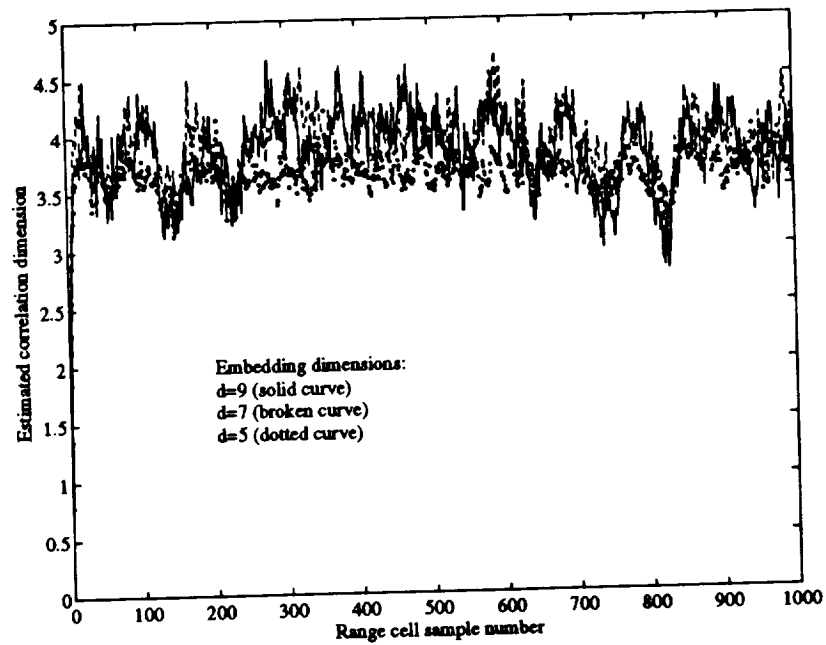


Figure 5.3 Calculated correlation dimension for ground clutter for the second test. Time delay is 5 samples. Embedding dimensions are 5, 7, and 9.

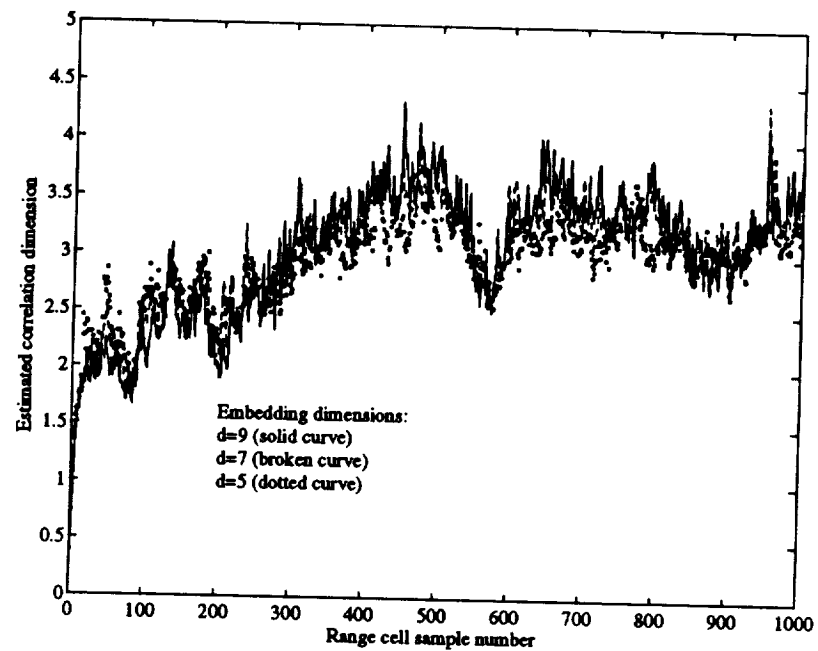


Figure 5.4 Calculated correlation dimension for weather for the second test. Time delay is 5 samples. Embedding dimensions are 5, 7, and 9.

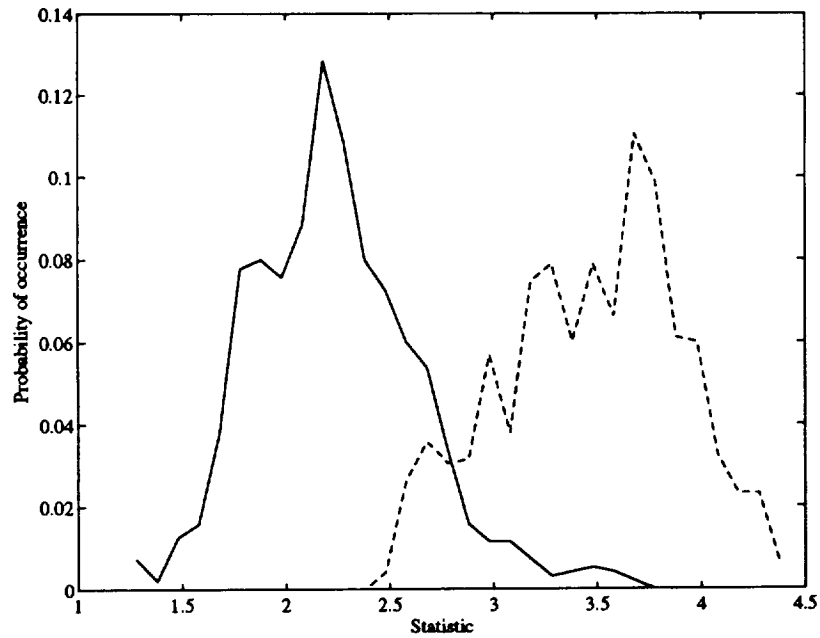


Figure 5.5 Density functions of correlation dimension values conditioned on ground returns (broken curve) and weather returns (solid curve) for the first test. Using a threshold of 2.78, it is possible to achieve a detection probability of 0.93 while maintaining a false alarm rate less than 0.10.

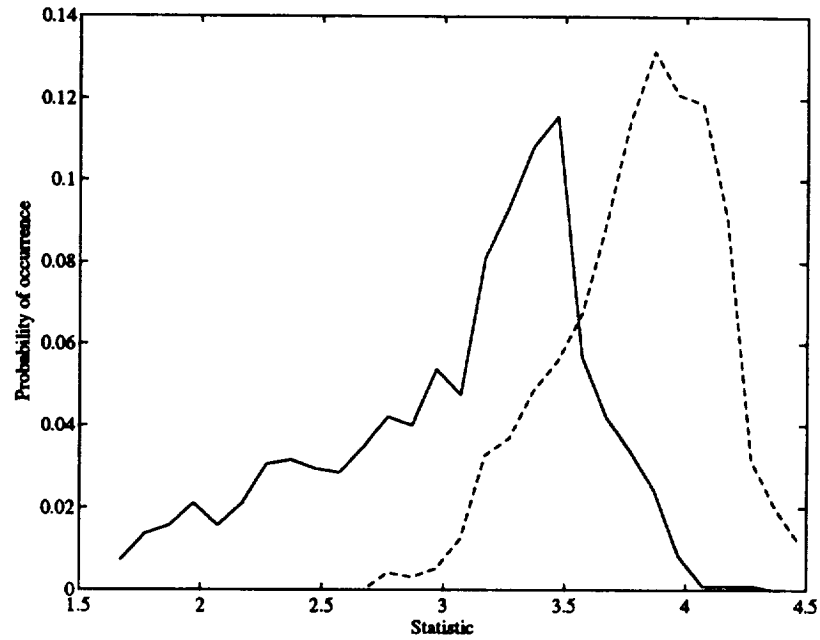


Figure 5.6 Density functions of correlation dimension values conditioned on ground returns (broken curve) and weather returns (solid curve) for the second test. Using a threshold of 3.7, it is possible to achieve a detection probability of 0.71 while maintaining a false alarm rate less than 0.10.



ground returns, as discussed earlier. The conclusion is that the correlation dimension is a promising discriminant but the threshold may have to be adjusted.

It appears that there is a significant difference in the correlation dimension between the weather and ground returns, with the weather return having a relatively small correlation dimension. Thus it seems reasonable that the correlation dimension may be effective in discriminating the source of radar returns, when those sources may be weather or the ground. However, much more detailed analyses will be required before these results can be presented as general.

## CHAPTER 6

### CONCLUSIONS

This work has been concerned with the estimation of the correlation dimension. The relationship between the correlation dimension and the structure of system attractors and estimation of the correlation dimension from a single time sequence have been discussed. In this context, the practical aspects of correlation dimension estimation have been emphasized. Further, as an application, it was shown that the correlation dimension can be used, at least in some cases, to distinguish between radar return signals of differing origins.

#### 6.1 Discussion of Results

##### 6.1.1 New Estimate of the Correlation Dimension

The practical estimation of the correlation dimension poses many problems; these include topics relating to the derivation of multidimensional dynamics from a single time sequence, as well as details of calculating the correlation integral from state vectors and calculating the correlation dimension from the correlation integral. Here effort was focused on the last of these topics.

The requirement that only the “small” range of the correlation integral be used in estimating the correlation dimension posed significant statistical problems. The traditional approach has been to find the slope of a line which fits the log-log plot of the correlation integral in a least squares sense. More recently, the Takens Estimator was developed, which provides an optimal estimate of the correlation dimension. However, the least squares method was shown to have theoretical problems, since the points of the correlation integral are not statistically independent, while the Takens Estimator does not provide significantly improved results.

As a practical matter, both the least squares and Takens Estimator require the *a priori* specification of a range of the correlation integral to consider. Calculations here demonstrated that the results obtained are strongly influenced by the range chosen. A new estimate for the correlation dimension was developed by choosing an error criterion for the least squares method which is: (1) more strongly affected by the linear portion of the log-log plot of the correlation integral than by the saturation region and (2) based on the increments in the correlation integral, eliminating the statistical dependence between points.

This new estimate thus removes both an implementation obstacle (point 1 above) and a theoretical obstacle (point 2 above). Results show that it provides an accurate estimate of the correlation dimension for several well known examples.

#### 6.1.2 Correlation Dimension as a Radar Reflector Discriminant

Perhaps the most common method for discriminating between radar returns from different reflectors is through the use of frequency selective filters. However, in applications such as airborne radar, in which the ground clutter cannot be simply removed by a notch filter, another method or measure is desirable. It is for such a situation, in which weather and clutter may occupy the same or very nearby frequency bands, that the use of the correlation dimension has been proposed as a discriminant.

The results presented for using the correlation dimension to discriminate between radar returns from weather and ground clutter indicate that it can be very effective for such a task. The work of Lorenz [1] demonstrated that weather is a chaotic system of relatively low dimension, and this is reflected in the correlation dimension estimates for the weather radar returns. These weather returns seem to imply a correlation dimension somewhere between 2.0 and 2.5. There seems to be little guidance for what to expect for the dimension of ground returns, however. Empirical evidence suggests a correlation dimension exceeding 2.5.

## 6.2 Conclusions and Future Work

Work in chaotic systems is relatively new. The idiosyncrasies of chaotic modelling would seem to render such work pointless, for the very nature of chaotic systems is that they are unpredictable. Indeed, if prediction were the only purpose, such work would be useless. However, in many cases it is beneficial to know simply the nature of the physical system underlying an observed time sequence, for example, is the sequence truly random, or is a chaotic model appropriate? In such a case, a *model* may not be required, but simply an idea of the *form* of the model.

If this is the case, *descriptors* of the model are appropriate. These descriptors should contain as much information about the system as possible and be readily derivable from the data at hand, which is typically a single time sequence. The correlation dimension fits these requirements for chaotic systems. While in this work many relevant topics have been presented, many questions are not answered conclusively. The time delay embedding procedure itself may contain many opportunities for introduction of errors. For example, evidence and arguments presented here seem to indicate that the choice of time delay for the time delay embedding based on mutual information considerations should be given great consideration.

One of the main concerns here has been in deriving the correlation dimension from the correlation integral. Even for this small detail, an ideal method has not been found. Here a method of estimation has been proposed which seems very promising, but is overly sensitive to noise in the data. However, insofar as precision in the estimate is often at least as important as accuracy, the new method is a great improvement over existing methods.

The second concern was using the correlation dimension to distinguish between weather radar returns and those from the ground. Some preliminary empirical evidence has been presented to support the premise that weather returns are characterized by low correlation dimension, while ground returns are significantly higher in

correlation dimension. Again the results are encouraging, but problems such as correlation dimension estimation from extremely short time sequences must be overcome.

## REFERENCES

1. Edward N. Lorenz. Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, 20:130–141, 1963.
2. Harold Froehling, J. P. Crutchfield, Doyne Farmer, N. H. Packard, and Rob Shaw. On determining the dimension of chaotic flows. *Physica D*, 3:605–617, 1981.
3. N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw. Geometry from a time series. *Physical Review Letters*, 45(9):712–716, 1980.
4. Peter Grassberger and Itamar Procaccia. Characterization of strange attractors. *Physical Review Letters*, 50(5):346–349, 1983.
5. Peter Grassberger and Itamar Procaccia. Measuring the strangeness of strange attractors. *Physica D*, 9:189–208, 1983.
6. J. Doyne Farmer, Edward Ott, and James A. Yorke. The dimension of chaotic attractors. *Physica D*, 7:153–180, 1983.
7. Floris Takens. Detecting strange attractors in turbulence. *Lecture Notes in Mathematics*, 898:366–381, 1981.
8. Floris Takens. On the numerical determination of the dimension of an attractor. *Lecture Notes in Mathematics*, 1125(11):99–106, 1985.
9. Andrew M. Fraser and Harry L. Swinney. Independent coordinates for strange attractors from mutual information. *Physical Review A*, 33(2):1134–1140, 1986.
10. James Theiler. *Quantifying Chaos: Practical Estimation of the Correlation Dimension*. PhD thesis, California Institute of Technology, Pasadena, California, 1988.
11. James Theiler. Estimating the fractal dimension of chaotic time series. *The Lincoln Laboratory Journal*, 3(1):63–86, 1990.
12. James Theiler. Estimating fractal dimension. *Journal of the Optical Society of America A*, 7(6):1055–1073, 1990.
13. Henry Leung and Simon Haykin. Is there a radar clutter attractor? *Applied Physics Letters*, 56(6):593–595, 1990.
14. Jill I. Butterfield. Fractal interpolation of radar signatures for detecting stationary targets in ground clutter. *IEEE AES Systems Magazine*, July 1991, pages 10–14.

15. Alan V. Oppenheim, Gregory W. Wornell, Steven H. Isabelle, and Kevin M. Cuomo. Signal processing in the context of chaotic signals. In *Proceedings of the 1992 International Conference on Acoustics, Speech, and Signal Processing*, San Francisco, CA, March 1992, pages IV – 117–120.
16. John J. Sidorowich. Modeling of chaotic time series for prediction, interpolation, and smoothing. In *Proceedings of the 1992 International Conference on Acoustics, Speech, and Signal Processing*, San Francisco, CA, March 1992, pages IV – 121–124.
17. Simon Haykin and Henry Leung. Model reconstruction of chaotic dynamics: First preliminary radar results. In *Proceedings of the 1992 International Conference on Acoustics, Speech, and Signal Processing*, San Francisco, CA, March 1992, pages IV – 125–128.
18. John C. Kirk, Jr. Using the fractal dimension (FD) to discriminate between targets and clutter. In *Proceedings of the 1993 IEEE National Radar Conference*, Boston, MA, April 1993, pages 76–78.
19. Kevin D. Barnett and Ernest G. Baxa, Jr. A study of the correlation dimension of weather and ground clutter returns in airborne radar. In *Proceedings of the American Meteorological Society's 26th International Conference on Radar Meteorology*, Norman, OK, May 1993, pages 247–248.
20. T. T. Fujita. The downburst, microburst and macroburst. SMRP Research Paper 210, University of Chicago, 1985.
21. Emedio M. Bracalente, W. R. Jones, and C. L. Britt. Airborne Doppler radar detection of low altitude windshear. In *Proceedings of AIAA Conference on Sensor and Measurement Techniques for Aeronautical Applications*, Atlanta, GA, September 1988.
22. Ernest G. Baxa, Jr. Airborne pulsed Doppler radar detection of low-altitude windshear — a signal processing problem. *Digital Signal Processing*, 1:186–197, 1991.
23. J.-P. Eckmann and D. Ruelle. Ergodic theory of chaos and strange attractors. *Reviews of Modern Physics*, 57(3, Part I):617–656, 1985.
24. Robert L. Devaney. *An Introduction to Chaotic Dynamical Systems*. The Benjamin/Cummings Publishing Co., Inc., Menlo Park, CA, 1986.
25. Henry D. I. Arbel. Chaotic signals and physical systems. In *Proceedings of the 1992 International Conference on Acoustics, Speech, and Signal Processing*, San Francisco, CA, March 1992, pages IV – 113–116.
26. Benoit B. Mandelbrot. *Fractals: Form, Chance, and Dimension*. W. H. Freeman, San Francisco, CA, 1977.

27. Benoit B. Mandelbrot. *The Fractal Geometry of Nature*. W. H. Freeman, San Francisco, CA, 1983.
28. Andrew M. Fraser. Reconstructing attractors from scalar time series: A comparison of singular system and redundancy criteria. *Physica D*, 34:391–404, 1989.
29. Robert G. Gallager. *Information Theory and Reliable Communication*. John Wiley and Sons, New York, 1968.







REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.				
1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE December 1993	3. REPORT TYPE AND DATES COVERED Contractor Report		
4. TITLE AND SUBTITLE On the Estimation of the Correlation Dimension and Its Application to Radar Reflector Discrimination		5. FUNDING NUMBERS G NAG1-928 WU 505-64-12-02		
6. AUTHOR(S) Kevin D. Barnett				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Radar Systems Laboratory Clemson University Clemson, SC 29634-0915		8. PERFORMING ORGANIZATION REPORT NUMBER TR-073093-3570P RSL Technical Report 17		
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) National Aeronautics and Space Administration Langley Research Center Hampton, VA 23681-0001		10. SPONSORING / MONITORING AGENCY REPORT NUMBER NASA CR-4564 DOT/FAA/RD-93/41		
11. SUPPLEMENTARY NOTES Langley Technical Monitor: Anne I. Mackenzie Final Report				
12a. DISTRIBUTION / AVAILABILITY STATEMENT Unclassified - Unlimited  Subject Category 33			12b. DISTRIBUTION CODE	
<p><del>Recently, system theorists</del> have recognized that low order systems of nonlinear differential equations can give rise to solutions which are neither periodic, constant, nor predictable in steady state, but which are nonetheless bounded and deterministic. This behavior, which was first described in the study of weather systems, has been termed "chaotic." Much study of chaotic systems has concentrated on analysis of the systems' phase space attractors. It has been recognized that invariant measures of the attractor possess inherent information about the system. One such measure is the dimension of the attractors. The dimension of a chaotic attractor has been shown to be noninteger, leading to the term "strange attractor"; the attractor is said to have a fractal structure. The correlation dimension has become one of the most popular measures of dimension. However, many problems have been identified in correlation dimension estimation from time sequences. The most common methods for obtaining the correlation dimension have been least squares curves fitting to find the slope of the correlation integral and the Takens Estimator. However, these estimates show unacceptable sensitivity to the upper limit on the distance chosen. Here, a new method is proposed which is shown to be rather insensitive to the upper limit and to perform in a very stable manner, at least in the absence of noise. The correlation dimension is also shown to be an effective discriminant in distinguishing between radar returns resulting from weather and those from the ground. The weather returns are shown to have a correlation dimension generally between 2.0 and 3.0, while ground returns have a correlation dimension exceeding 3.0.</p>				
14. SUBJECT TERMS weather radar, chaotic systems, strange attractor, correlation dimension estimation, target discrimination			15. NUMBER OF PAGES 76	
			16. PRICE CODE A05	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT	

